

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2002, 07:13:24 ; Search time 3144 Seconds

(without alignments)  
8034.744 Million cell updates/sec

Title: US-09-921-994-1

Perfect score: 868  
Sequence: 1 ataagaatgcgcgcgcgatga.....ctggcgcaagtagactagtc 868

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	868	100.0	868	6	AX377958	AX377958 Sequence
2	686.2	79.1	1161	6	AR050784	AR050784 Sequence
3	686.2	79.1	1195	9	BC015364	BC015364 Homo sapi
4	683	78.7	1161	6	AR003351	AR003351 Sequence
5	683	78.7	1161	9	HUMERI3X	L08187 Human cyto
6	354	40.8	1131	10	BC008209	BC008209 Mus muscu
7	354	40.8	1161	10	AF013114	AF013114 Mus muscu
8	205.2	23.6	445	6	AX392987	AX392987 Sequence
9	182.6	21.0	43297	9	AC005578	AC005578 Homo sapi
10	172.2	19.8	179607	2	AC027272	AC027272 Homo sapi
11	102.8	11.8	1116	11	G54090	G54090 902 Human H
12	97	11.2	193355	2	AC073737	AC073737 Mus muscu
13	90.2	10.4	176359	2	AC118074	AC118074 Rattus no
14	74.8	8.6	1332	6	AX399177	AX399177 Sequence
15	74.8	8.6	1332	10	S54212	S54212 ciliary neu
16	74.8	8.6	1975	10	AF068615	AF068615 Mus muscu
17	73.2	8.4	125020	9	AF429315	AF429315 Homo sapi
18	69	7.9	1566	6	AX399178	AX399178 Sequence
19	69	7.9	1566	9	HUMCNTER	M73238 Human cilia
20	69	7.9	1591	6	AR066232	AR066232 Sequence
21	69	7.9	1591	6	AR070290	AR070290 Sequence
22	69	7.9	1591	6	AR074863	AR074863 Sequence
23	69	7.9	1591	6	I12552	I12552 Sequence 1
24	69	7.9	1591	6	I56052	I56052 Sequence 1
25	69	7.9	1692	6	AX205060	AX205060 Sequence
26	69	7.9	1923	9	BC001492	BC001492 Homo sapi
27	69	7.9	2020	6	AX281744	AX281744 Sequence
28	67.4	7.8	1089	5	GGGPARA	U29245 Gallus gall
29	65.8	7.6	1248	5	GGGPARA	Z48168 G.gallus mr
30	54.8	6.3	125020	9	AF429315	AF429315 Homo sapi
31	54.2	6.2	109523	2	AC111892	AC111892 Rattus no
32	53.8	6.2	10732	6	E32986	E32986 Gene encodi
33	52	6.0	114062	2	OSJN00267	AL731619 Oryza sat
34	51	5.9	3060	1	SAUREVTRAN	M86352 Stigmatella
35	50.8	5.9	159851	2	AP003571	AP003571 Oryza sat
36	50.8	5.9	209887	2	AC079424	AC079424 Mus muscu
37	50.6	5.8	298166	2	AC087563	AC087563 Homo sapi
38	50.4	5.8	61958	2	AC121467	AC121467 Rattus no
39	49.8	5.7	249262	2	AC079430	AC079430 Mus muscu
40	49.8	5.7	303091	2	AC084799	AC084799 Mus muscu
41	49	5.6	7218	6	I66494	I66494 Sequence 14
42	49	5.6	158880	2	AC128264	AC128264 Rattus no
43	49	5.6	300695	2	AC079431	AC079431 Mus muscu
44	48.8	5.6	110000	2	AC098456_1	Continuation (2 of
45	48.4	5.6	130540	2	AC079417	AC079417 Mus muscu

ALIGNMENTS

RESULT 1	AX377958	868 bp	DNA	linear	PAT 18-MAR-2002
LOCUS	AX377958	Sequence 1 from Patent WO0212282.			
DEFINITION	AX377958				
ACCESSION	AX377958				
VERSION	AX377958.1	GI:19574012			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1	Bowman,M.R.			
AUTHORS	Novel ebi-3-alt protein and nucleic acid molecules and uses				
TITLE	therefor				

JOURNAL	Patent: WO 0212282-A 1 14-FEB-2002;
GENETICS	INSTITUTE, INC. (US)
FEATURES	Location/Qualifiers
source	1. .868
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	147 a 325 c 239 g 157
ORIGIN	

Query Match	100.0%;	Score 868;	DB 6;	Length 868;
Best Local Similarity	100.0%;	Pred. No. 5.1e-165;		
Matches 868; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0.

QY		1	ATAGAATGCGGCCGCATGACCCCCCAGACTTCTCTGGCCCTGTCTCTGTGGCCAGCTG	60
Db		1	ATAGAATGCGGCCGCATGACCCCCCAGACTTCTCTGGCCCTGTCTCTGTGGCCAGCTG	60
QY		61	CCCCCCCCCAGTGGAAGAAAGGGCCCCCAGCAGACTCTGACACTGCCCCGGGTGCAATG	120
Db		61	CCCCCCCCCAGTGGAAGAAAGGGCCCCCAGCAGACTCTGACACTGCCCCGGGTGCAATG	120
QY		121	CCGAGCCTTCGCGTACCAGATGCGCGGTGATTGCTCTGAGACCTGCGCATGACCCCGC	180
Db		121	CCGAGCCTTCGCGTACCAGATGCGCGGTGATTGCTCTGAGACCTGCGCATGACCCCGC	180
QY		181	AGCTTCTCCTGCGCCTTTGTCTCTGCGCCAGCTGCCCCCGCCTGCGGTGAAGAAAGGC	240
Db		181	AGCTTCTCCTGCGCCTTTGTCTCTGCGCGCACGTGCCCGCCTGCGGTGAAGAAAGGC	240
QY		241	CCCCAGCAGCTTGACACTGCCCCCGGTGCAATGCCAGCCTCGGTACCCGATGCGCG	300
Db		241	CCCCAGCAGCTTGACACTGCCCCCGGTGCAATGCCAGCCTCGGTACCCGATGCGCG	300
QY		301	TGGATTGCTCTTGACCCCTGCGCGCTGCTCCAACCTCCACCAACCCCCTGTCTTCAATTG	360
Db		301	TGGATTGCTCTTGACCCCTGCGCGCTGCTCCAACCTCCACCAACCCCCTGTCTTCAATTG	360
QY		361	CCACGTACAGGCTCGGCATGGCTGCCCGGGGCCACAGCTGGCCTTGCTTGACAGCAGCGC	420
Db		361	CCACGTACAGGCTCGGCATGGCTGCCCGGGGCCACAGCTGGCCTTGCTTGACAGCAGCGC	420
QY		421	CAACGTCCACAGCTGCACCATCACGGATGTCCAGCTGTTCTCCATGGCTCCCTACGTGC	480
Db		421	CAACGTCCACAGCTGCACCATCACGGATGTCCAGCTGTTCTCCATGGCTCCCTACGTGC	480
QY		481	TCAATGTACCCGCGTCCGCCCTGGGGCTCCAGCAGCAGCTTCGTTCATTAACAG	540
Db		481	TCAATGTACCCGCGTCCGCCCTGGGGCTCCAGCAGCAGCTTCGTTCATTAACAG	540
QY		541	AGCACATCATCAAGCCCCGACCCTCCAGAAGGCGTGCCTAAGCCCCCTCGCTGAGCGCC	600
Db		541	AGCACATCATCAAGCCCCGACCCTCCAGAAGGCGTGCCTAAGCCCCCTCGCTGAGCGCC	600
QY		601	AGCTACAGGTGCA GTGGGAGCCTCCCGGCTCTGGCCCTTCCAGAGATCTTCTCACTGA	660
Db		601	AGCTACAGGTGCA GTGGGAGCCTCCCGGCTCTGGCCCTTCCAGAGATCTTCTCACTGA	660
QY		661	AGTACTGGATCCGTTACAAGCGTCAGGAGACTGCGGCTTCCACCGGGTGGGGCCCA TTG	720
Db		661	AGTACTGGATCCGTTACAAGCGTCAGGAGACTGCGGCTTCCACCGGGTGGGGCCCA TTG	720
QY		721	AAGCCACGCTCTTCATCTCAGGGCGTGTGCGCCCCGAGCCAGGTACTACGTCCAAGTGG	780
Db		721	AAGCCACGCTCTTCATCTCAGGGCGTGTGCGCCCCGAGCCAGGTACTACGTCCAAGTGG	780
QY		781	CGGCTCAGGACTTCACAGACTACGGGGAAGTGA GTGACTGAGTCTCCCCCGCACTGCCA	840
Db		781	CGGCTCAGGACTTCACAGACTACGGGGAAGTGA GTGACTGAGTCTCCCCCGCACTGCCA	840
QY		841	CAATGAGCCTGGGCAAGTAGACTAGTTC	868
Db		841	CAATGAGCCTGGGCAAGTAGACTAGTTC	868
Db		841	CAATGAGCCTGGGCAAGTAGACTAGTTC	868

RESULT	2
AR050784	
LOCUS	AR050784
DEFINITION	Sequence 1 from patent US 5830451.
ACCESSION	AR050784
VERSION	AR050784.1 GI:5974148
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified. 1 (bases 1 to 1161) Devergne,O. and Kieff,E.D. Haematopoietic cytokine Epstein Barr virus-induced protein Patent: US 5830451-A 1 03-NOV-1998;
AUTHORS	
TITLE	
JOURNAL	
FEATURES	Location/Qualifiers source 1..1161
BASE COUNT	239 a /organism="unknown" 378 c 304 g 240 t
ORIGIN	

Query Match	79.18;	Score 686.2;	DB 6;	Length 1161;
Best Local Similarity	99.68;	Pred. No. 2.1e-128;		
Matches 688; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	170	CATGACCCCGCAGCTTCTCTCTGGCCCTGTCTCTCTGGGCCAGCTGCCCCCTGGCGGTGG	229
Db	13	CATGACCCCGCAGCTTCTCTCTGGCCCTGTCTCTCTGGGCCAGCTGCCCCCTGCAGTGG	72
QY	230	AAGGAAGGGCCCCCAGCAGCTCTGACACTGCCCCGGGTGCATATGCCGAGCTCTCGGTA	289
Db	73	AAGGAAGGGCCCCCAGCAGCTCTGACACTGCCCCGGGTGCATATGCCGAGCTCTCGGTA	132
QY	290	CCCGATCGCCGTGATGTGCTCTGTGACCCCTGCGCCTGTCTCCAACTCCACAGCCCGT	349
Db	133	CCCGATCGCCGTGATGTGCTCTGTGACCCCTGCGCCTGTCTCCAACTCCACAGCCCGT	192
QY	350	GTCCTTCATTGCGCAGTACAGGCTCGGCAATGGCTGGCCCGGGCCACAGCTGAGCTTGCT	409
Db	193	GTCCTTCATTGCGCAGCTACAGGCTCGGCAATGGCTGGCCCGGGCCACAGCTGAGCTTGCT	252
QY	410	GCAGCAGACGCCCAACGCTCCACCAGCTGCAACCATCACGGATGTCCAGCTGTCTCCATGGC	469
Db	253	GCAGCAGACGCCCAACGCTCCACCAGCTGCAACCATCACGGATGTCCAGCTGTCTCCATGGC	312
QY	470	TCCCTACGTGCTCAATGTCAACCGCCGTCGCCCCCTGGGGCTCCAGCAGCAGCTTCGTGCC	529
Db	313	TCCCTACGTGCTCAATGTCAACCGCCGTCGCCCCCTGGGGCTCCAGCAGCAGCTTCGTGCC	372
QY	530	TTTCATTACAGAGCAGCATCATCAAGCCCCGACCTCCAGAAAGGCGTGCCTAAGCCCCCT	589
Db	373	TTTCATTACAGAGCAGCATCATCAAGCCCCGACCTCCAGAAAGGCGTGCCTAAGCCCCCT	432
QY	590	CGCTGAGCGCCAGCTACAGGTGCAGTGGGAGCCTCCCGGCTCTGGCCCTTCCACAGAT	649
Db	433	CGCTGAGCGCCAGCTACAGGTGCAGTGGGAGCCTCCCGGCTCTGGCCCTTCCACAGAT	492
QY	650	CTTCTCACTGAAGTACTGGATCCGTTACAGCGTCAAGGAGCTGCGCGCTTCCACCGGGT	709
Db	493	CTTCTCACTGAAGTACTGGATCCGTTACAGCGTCAAGGAGCTGCGCGCTTCCACCGGGT	552
QY	710	GGGGCCCATTTGAAGCCAGCTCTTCATCTCAAGGCTGTGCGGCCCCGAGCCAGGTACTA	769
Db	553	GGGGCCCATTTGAAGCCAGCTCTTCATCTCAAGGCTGTGCGGCCCCGAGCCAGGTACTA	612
QY	770	CGTCCAAGTGGCGGCTCAGGACCTCACAGACTACGGGAACCTGAGTGAAGTGAAGTCC	829
Db	613	CGTCCAAGTGGCGGCTCAGGACCTCACAGACTACGGGAACCTGAGTGAAGTGAAGTCC	672
QY	830	CGCCACTGCCACATGAGCCTGGGCAAGTAG	860
Db	673	CGCCACTGCCACATGAGCCTGGGCAAGTAG	703

RESULT 3  
LOCUS BC015364 1195 bp mRNA linear PRI 04-OCT-2001  
DEFINITION Homo sapiens, Epstein-Barr virus induced gene 3, clone MGC:21515  
IMAGE:3897136, mRNA, complete cds.  
ACCESSION BC015364  
VERSION BC015364.1 GI:15929884  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1195)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: http://www-shgc.stanford.edu  
Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: Mgc clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
Series: IRAK Plate: 22 Row: f Column: 18  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.  
FEATURES  
source location/Qualifiers  
1..1195  
/organism="Homo sapiens"  
/db\_xref="LocusID:10148"  
/db\_xref="taxon:9606"  
/clone="MGC:21515 IMAGE:3897136"  
/tissue\_type="Pancreas, epithelioid carcinoma"  
/clone\_1lb="NIH\_MGC\_70"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
54..743  
/codon\_start=1  
/product="Epstein-Barr virus induced gene 3"  
/protein\_id="AAH15364.1"  
/db\_xref="GI:15929885"  
/translation="MTPQLLALVLMASCPSCGKRPALTLPRVQCRASRYPIAV  
DCSWTLPPAPNSTSPVSEIATYRLGMAARGHSMPCIQOTPTSTCTIDVOLFSMAPY  
VLNVTAVHPMGSSSSSFVPFITEHIKPDPEGVRLSPLAERQLQVQWEPGSPFPEI  
FSLKYWIRYKROGAARFHRVGPTEATSFILRAVRPRARYVQVAAQDLTDYGLSDMS  
LPATATMSLGR"  
CDS  
BASE COUNT 247 a 395 c 311 g 242 t  
ORIGIN  
Query Match 79.1%; Score 686.2; DB 9; Length 1195;  
Best Local Similarity 99.6%; Pred. No. 2e-128;  
Matches 688; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 170 CATGACCCCGCAGCTTCTCCTGGCCCTTGTCCTGTGGCCAGCTGCCCGCCGCGGTGG 229  
|||||  
Db 53 CATGACCCCGCAGCTTCTCCTGGCCCTTGTCCTGTGGCCAGCTGCCCGCCGCGAGTGG 112  
QY 230 AAGGAAGGGCCCCCAGCAGCTGTGACACTGCCCCGGGTGCAATGCCGAGCCTCTCGGTA 289  
|||||  
Db 113 AAGGAAGGGCCCCCAGCAGCTGTGACACTGCCCCGGGTGCAATGCCGAGCCTCTCGGTA 172

QY 290 CCCGATGCGCGTGGATGTGCTCCTGGACCCCTGCGCCCTGCTCCAAACTCCACGACCCCGT 349  
|||||  
Db 173 CCCGATGCGCGTGGATGTGCTCCTGGACCCCTGCGCCCTGCTCCAAACTCCACGACCCCGT 232  
QY 350 GTCCCTCATTTGCCACGTTACAGGCTCGGCAATGGCTGCCCCGGGGCCACAGCTGGCCTTGCCCT 409  
|||||  
Db 233 GTCCCTCATTTGCCACGTTACAGGCTCGGCAATGGCTGCCCCGGGGCCACAGCTGGCCTTGCCCT 292  
QY 410 GCAGCAGACGCCCAACGCTCCACCAGCTGCACCATCAGCGATGTCCAGCTGTCTCCATGGC 469  
|||||  
Db 293 GCAGCAGACGCCCAACGCTCCACCAGCTGCACCATCAGCGATGTCCAGCTGTCTCCATGGC 352  
QY 470 TCCCTACGTGCTCAATGTCAACCGCCGTCGCCCTGGGGCTCCAGCAGCAGCTTCTGTCGC 529  
|||||  
Db 353 TCCCTACGTGCTCAATGTCAACCGCCGTCGCCCTGGGGCTCCAGCAGCAGCTTCTGTCGC 412  
QY 530 TTTCAATAACAGACACATCATCAAGCCCGACCCCTCCAGAAGCGGTGCGCCTAAGCCCCCT 589  
|||||  
Db 413 TTTCAATAACAGACACATCATCAAGCCCGACCCCTCCAGAAGCGGTGCGCCTAAGCCCCCT 472  
QY 590 CGCTGAGCGCCAGCTACAGGTGTCAGTGGGAGCCTCCCGGGTCTGTGCCCCCTTCCACAGAGAT 649  
|||||  
Db 473 CGCTGAGCGCCAGCTACAGGTGTCAGTGGGAGCCTCCCGGGTCTGTGCCCCCTTCCACAGAGAT 532  
QY 650 CTTCTCACTGAAGTACTGTGATCCGTTTACAAGCGTCAAGGAGCTGCGCGCTTCCACCGGGT 709  
|||||  
Db 533 CTTCTCACTGAAGTACTGTGATCCGTTTACAAGCGTCAAGGAGCTGCGCGCTTCCACCGGGT 592  
QY 710 GGGGCCCATTTGAAGCCACAGTCTCTTCAATCCTCAAGGGCTGTGCGGCCCGAGCCAGGTACTA 769  
|||||  
Db 593 GGGGCCCATTTGAAGCCACAGTCTCTTCAATCCTCAAGGGCTGTGCGGCCCGAGCCAGGTACTA 652  
QY 770 CGTCCAAGTGGCGGCTCAGGACCTCAGACTACAGGGGAACGTGAGTGTGAGTCTCCC 829  
|||||  
Db 653 CGTCCAAGTGGCGGCTCAGGACCTCAGACTACAGGGGAACGTGAGTGTGAGTCTCCC 712  
QY 830 CGCCACTGCGACAATGAGCCTGGGCAAGTAG 860  
|||||  
Db 713 CGCCACTGCGACAATGAGCCTGGGCAAGTAG 743  
RESULT 4  
LOCUS AR003351 1161 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 5 from patent US 5744301.  
ACCESSION AR003351  
VERSION AR003351.1 GI:3964610  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1161)  
AUTHORS Birkenbach, M. and Kieff, E.  
TITLE Methods of detection of Epstein Barr virus induced genes expressed  
in the placenta  
JOURNAL Patent: US 5744301-A 5 28-APR-1998;  
FEATURES location/Qualifiers  
source 1..1161  
BASE COUNT 239 a 378 c 304 g 240 t  
ORIGIN  
Query Match 78.7%; Score 683; DB 6; Length 1161;  
Best Local Similarity 99.3%; Pred. No. 9e-128;  
Matches 686; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 170 CATGACCCCGCAGCTTCTCCTGGCCCTTGTCCTGTGGCCAGCTGCCCGCCCTGCGGTGG 229  
|||||  
Db 13 CATGACCCCGCAGCTTCTCCTGGCCCTTGTCCTGTGGCCAGCTGCCCGCCCTGCAAGTGG 72  
QY 230 AAGGAAGGGCCCCCAGCAGCTGTGACACTGCCCCGGGTGCAATGCCGAGCCTCTCGGTA 289  
|||||

Db 73 AAGGAAAGGGCCCCCAGCAGCTCTGACACTGCCCCGGGTGCAATGCCGAGCCTCTCGGTA 132

QY 290 CCCGATCGCGGTGATGCTCTCTGGACCCCTGCCGCTGCTCCAAACTCCACGAGCCCGT 349

Db 133 CCCGATCGCGGTGATGCTCTCTGGACCCCTGCCGCTGCTCCAAACTCCACGAGCCCGT 192

QY 350 GTCCCTCATTTGCCACGTACAGGCTCGGCATGGCTGCCGGGGCCACAGCTGGCCTTGCCCT 409

Db 193 GTCCCTCATTTGCCACGTACAGGCTCGGCATGGCTGCCGGGGCCACAGCTGGCCTTGCCCT 252

QY 410 GCAGCAGAGCCCAACGCTCCACCAAGCTGCACCATCAGGATGTCCAGCTGTCTCCATGGC 469

Db 253 GCAGCAGAGCCCAACGCTCCACCAAGCTGCACCATCAGGATGTCCAGCTGTCTCCATGGC 312

QY 470 TCCCTACGTGCTCAATGTACACCGCCGCTCCACCCCTGGGGGCTCCAGCAGCAGCTTCGTGCC 529

Db 313 TCCCTACGTGCTCAATGTACACCGCCGCTCCACCCCTGGGGGCTCCAGCAGCAGCTTCGTGCC 372

QY 530 TTTTCATTAACAGAGCACAATCATCAAGCCCGACCCCTCCAGAGAGGCGTGGCCCTTAAGCCCCCT 589

Db 373 TTTTCATTAACAGAGCACAATCATCAAGCCCGACCCCTCCAGAGAGGCGTGGCCCTTAAGCCCCCT 432

QY 590 CGCTGAGCGCCAGCTACAGGTGCTAGTGGAGCCTCCCGGCTCTGCGCCCTTCCAGAGAT 649

Db 433 CGCTGAGCGCCAGCTACAGGTGCTAGTGGAGCCTCCCGGCTCTGCGCCCTTCCAGAGAT 492

QY 650 CTTCCTACTGAAGTACTGATCCGTTTACAAGCGTCAAGGAGGCTGCGCTTCCAGCCGGGT 709

Db 493 CTTCCTACTGAAGTACTGATCCGTTTACAAGCGTCAAGGAGGCTGCGCTTCCAGCCGGGT 552

QY 710 GGGGCCCCATTGAAAGCCACGCTCTTCATCTCTCAGGGCTGTCCGGCCCCCGAGCCAGTACTA 769

Db 553 GGGGCCCCATTGAAAGCCACGCTCTTCATCTCTCAGGGCTGTCCGGCCCCCGAGCCAGTACTA 612

QY 770 CGTCCAAGTGGCGGCTCAGGACCTTCACAGACTACGGGGAAGTGAAGTGAAGTCTCCC 829

Db 613 CGTCCAAGTGGCGGCTCAGGACCTTCACAGACTACGGGGAAGTGAAGTGAAGTCTCCC 672

QY 830 CGCCACTGCCACAATGAGCTGGGCAAGTAG 860

Db 673 CGCCACTGCCACAATGAGCTGGGCAAGTAG 703

RESULT 5

HUMEB13X 1161 bp mRNA linear PRI 01-APR-1996

LOCUS Human cytokine receptor (EBI3) mRNA, complete cds.

DEFINITION L08187

ACCESSION L08187.1 GI:632973

VERSION cytokine receptor.

KEYWORDS Homo sapiens cDNA to mRNA.

SOURCE Homo sapiens

ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Devergne, O., Hummel, M., Koepfen, H., Le Beau, M.M., Nathanson, E.C., Kieff, E. and Birkenbach, M. A novel interleukin-12 p40-related protein induced by latent Epstein-Barr virus infection in B lymphocytes J. Virol. 70 (2), 1143-1153 (1996)

JOURNAL MEDLINE 96135230

PUBMED 8551575

COMMENT On Jan 24, 1995 this sequence version replaced gi:181921.

FEATURES

source

1. 1161

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_line="BL41/B95-8"

/cell\_type="B lymphocyte, EBV-converted Burkitt lymphoma"

1. 1161

/gene="EBI3"

14. 703

/gene="EBI3"

/codon\_start=1

/product="cytokine receptor"

/protein\_id="AA93193.1"

/db\_xref="GI:632974"

/translation="MTPQLLLALVLMASCPGSGRKGPAAALLPRVQCRASRYPIAV DCSWTLPPAPNSTSPVSFIATYRLGMAARHSHWPCLOQTPTSTCTITDVQLESMAPY VLNTAVHPNWSSSSFVPIETHEIIRKPDPEGVRLSPLAERHYOVQWEPGSGWPFPEI FSLKWIRYRKQGAARFHRVGPTEATSFILRAVRPRARYVOVAAQDLTDYGLSDWS LPATATMSLGK"

14. 73

sig\_peptide

14. 73

/gene="EBI3"

74. 700

mat\_peptide

/gene="EBI3"

/product="cytokine receptor"

BASE COUNT 239 a 378 c 304 g 240 t

ORIGIN

Query Match 78.7%; Score 683; DB 9; Length 1161;

Best Local Similarity 99.3%; Pred. No. 9e-128;

Matches 686; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 170 CATGACCCCGCAGCTTCTCTGACCCTTGCTCTGTGGCCAGCTGCCCGCCCTGCGGTG 229

Db 13 CATGACCCCGCAGCTTCTCTGACCCTTGCTCTGTGGCCAGCTGCCCGCCCTGCGGTG 72

QY 230 AAGGAAAGGGCCCCCAGCAGCTCTGACACTGCCCCGGGTGCAATGCCGAGCCTTCGGTA 289

Db 73 AAGGAAAGGGCCCCCAGCAGCTCTGACACTGCCCCGGGTGCAATGCCGAGCCTTCGGTA 132

QY 290 CCCGATCGCGGTGATGCTCTCTGACCCCTGCCGCTCTCCAAACTCCACGAGCCCGT 349

Db 133 CCCGATCGCGGTGATGCTCTCTGACCCCTGCCGCTCTCCAAACTCCACGAGCCCGT 192

QY 350 GTCCCTCATTTGCCACGTACAGGCTCGGATGGCTGCCCGGGGCCACAGCTGGCCTTGCCCT 409

Db 193 GTCCCTCATTTGCCACGTACAGGCTCGGATGGCTGCCCGGGGCCACAGCTGGCCTTGCCCT 252

QY 410 GCAGCAGAGCCCAACGCTCCACCAAGCTGACCATCAGGATGTCCAGCTGTCTCCATGGC 469

Db 253 GCAGCAGAGCCCAACGCTCCACCAAGCTGACCATCAGGATGTCCAGCTGTCTCCATGGC 312

QY 470 TCCCTACGTGCTCAATGTACACCGCCGCTCCCGCCCTGGGGGCTCCAGCAGCAGCTTCGTGCC 529

Db 313 TCCCTACGTGCTCAATGTACACCGCCGCTCCACCCCTGGGGGCTCCAGCAGCAGCTTCGTGCC 372

QY 530 TTTTCATTAACAGAGCACAATCATCAAGCCCGACCCCTCCAGAAAGGCGTGGCCTTAAGCCCCCT 589

Db 373 TTTTCATTAACAGAGCACAATCATCAAGCCCGACCCCTCCAGAAAGGCGTGGCCTTAAGCCCCCT 432

QY 590 CGCTGAGCGCCAGCTACAGGTGAGTGGAGGCTCCCGGCTCTGCGCCCTTCCACAGAGAT 649

Db 433 CGCTGAGCGCCAGCTACAGGTGAGTGGAGGCTCCCGGCTCTGCGCCCTTCCACAGAGAT 492

QY 650 CTTCCTACTGAAGTACTGATCCGTTTACAAGCGTCAAGGAGGCTGCGCGCTTCCACCGGCT 709

Db 493 CTTCCTACTGAAGTACTGATCCGTTTACAAGCGTCAAGGAGGCTGCGCGCTTCCACCGGCT 552

QY 710 GGGGCCCCATTGAAAGCCACGCTCTTCATCTCAGGGCTGTGGCGGCCCGAGCCAGGTA 769

Db 553 GGGGCCCCATTGAAAGCCACGCTCTTCATCTCAGGGCTGTGGCGGCCCGAGCCAGGTA 612

QY 770 CGTCCAAGTGGCGGCTCAGGACCTTCACAGACTACGGGGAAGTGAAGTGAAGTCTCCC 829

Db 613 CGTCCAAGTGGCGGCTCAGGACCTTCACAGACTACGGGGAAGTGAAGTGAAGTCTCCC 672

QY 830 CGCCACTGCCACAATGAGCCTGGGCAAGTAG 860

Db 673 CGCCACTGCCACAATGAGCCTGGGCAAGTAG 703

RESULT 6

BC008209 1131 bp mRNA linear ROD 07-AUG-2002

LOCUS BC008209 1131 bp mRNA linear ROD 07-AUG-2002

gene

CDS

/gene="EBI3"

14. 703

/gene="EBI3"



DEFINITION Mus musculus, Epstein-Barr virus induced gene 3, clone MGC:11457  
IMAGE:3154957, mRNA, complete cds.  
ACCESSION BC008209  
VERSION BC008209.1 GI:14198294  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1131)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: amg@bcm.tmc.edu  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAK Plate: 16 Row: c Column: 19  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657051.  
FEATURES  
source  
1. 1131  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="129,C57BL/6J,FVB/N"  
/clone="MGC:11457 IMAGE:3154957"  
/tissue\_type="Mammary tumor. Brca1-/f1; MMTV-Cre model. 10 months old, gross tissue."  
/clone\_1ib="NCI\_CGAP\_Mam3"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
188. 874  
/codon\_start=1  
/product="Epstein-Barr virus induced gene 3"  
/protein\_id="AAH08209.1"  
/db\_xref="GI:14198295"  
/db\_xref="LocustID:50498"  
/translation="MSKLFSLALMASRSPGYTETALVALSQPRVQCHASRYPAVD CSWTPLOAPNSTRSTSFIATYRLGVATQQSQPCLQSRPQASRCTIPDVHLFSTVPM LNTAVHPGGASSSLAFAVARILIKPDPEGVRLRTAGQRLQVLWHPASWPFEDIFS LKYLRLYRRRGASHFRQVGPIEATTTFTLRNSKPHAKYCIQVSAQDLTDYGKPSDWSLP GQVESAPHKP"  
BASE COUNT 242 a 374 c 268 g 247 t  
ORIGIN  
Query Match 40.8%; Score 354; DB 10; Length 1131;  
Best Local Similarity 66.5%; Pred. No. 1.7e-61;  
Matches 539; Conservative 0; Mismatches 265; Indels 6; Gaps 2;  
QY 22 CCGGCAGCTTCTCGTGGCCCTTGCTCTTGCGCAGCTGCCGCCCGCAGTGAAGAA 81  
DB 39 CCATCATCATCTCTGCCCATACACTGACACTGAGCCACACTGGGAGGCTCTTCCC 98  
QY 82 AGGGCCCCCAGCAGCTGTGACACTGCCCCGGGTGCAATGCCAGCCTCTGCGTACCCGAT 141  
DB 99 TGGGGCCAGGTTCCCTGTGTGAGTCCCTGTCTTCAACCCTCTCTGTGATGGGTCACTAA 158

QY 142 CGCCGTGAGTTGCTCTGTGACCCCTGCCGATGACCCCGCAGCTTCTCTGCGCCCTTGTCC 201  
DB 159 CTCGATCCAGGAAGACAGAGCCACAGAGCATGTCCAAAGCTGCTCTTCTGTCACTTGCCC 218  
QY 202 TCTGGCCAGCTGCCCCGCTTGGGTGGAAGAAAGGGCCCCCAGCAGCTCTGACACTGC 261  
DB 219 TCTGGGCCAGCCGCTCCCT--GGTTACACTGAACACAGCTCTCGTGCTTAAGCCAGC 275  
QY 262 CCGGGGTGCAATGCCGAGCCTCTCGGTACCCGATCGCCCGTGATTTGCTCTGACCCCTGC 321  
DB 276 CCAGAGTGCATGCCATGCTTCTCGGTATCCCGTGCGCGTGAGCTGCTCTCGACTCTC 335  
QY 322 CGCCTGCTCCAACTCCACCCAGCCCGTGTCTCTCATTTGCCACGTACAGGCTCGGATGG 381  
DB 336 TCCAGGCTCCCAACTCCACAGATCCAGTCCCTTCAATTGCGCACTTACAGGCTCGGTGG 395  
QY 382 CTGCCCCGGGCCACAGCTGGCCTTGCTGCTGACGACAGCCCAAGCTCCACAGCTGACCA 441  
DB 396 CCACCAGCAGCAGAGCCAGCCCTGCTTACCAACGAGAGCCCGCCTCCGATGACCA 455  
QY 442 TCACGATGTCCAGCTGTCTTCCATGGCTCCCTACGTGTCAATGTCAACGCCGCTCGCC 501  
DB 456 TCCCCGACGTGCACCTGTCTTCCACGGTGCCCTACATGTAATGTCACTGCAGTGCACC 515  
QY 502 CCTGGGCTCCAGCAGCAGCTTCGTGCTTTCATTAACAGACACATCATCAAGCCCGACC 561  
DB 516 CAGCGCGGCCACAGCAGACGCTTCTAGCCTTGTGGCTGAGCGAATCATCAAGCCGAGC 575  
QY 562 CTCAGAGAGGCGGTGGCCTTAAGCCCCCTCGCTGAGCCGCCACGTACAGGTGCAGTGGAGC 621  
DB 576 CTCGGAGAGCGGTGGCCTGCGCCACAGCGGGACAGCGC--CTGCAGGTGCTCTGCGATC 632  
QY 622 CTCGCCGGTCTTGCCCCCTTCCACAGATCTTCTACACTGAAGTACTGGATCCGTTACAAGC 681  
DB 633 CCCCTGCTTCTTGCCCCCTTCCGACATCTTCTCTCAAGTACCGACTCCGCTACCGGC 692  
QY 682 GTCAGGAGCTGCGGCTTCCACCGGGTGGGCCCATTTGAAGCCACGTCCTTCATCCTCA 741  
DB 693 GCCGAGAGCCTTCTCACTTCCGCAAGGTGGAGCCCATTTGAAGCCACGACTTTCACCTCA 752  
QY 742 GGGCTGTGGCCCCCGAGCCAGTACTACGTCCAAAGTGGCGGCTCAGACCTCAGAGACT 801  
DB 753 GGAAGTCGAACCCCATGCGCAAGTATTCATCCAGGTGTCAAGCTCAGACCTCAGAGATT 812  
QY 802 ACGGGAAGTGAAGTGAAGTGTCTCCCG 831  
DB 813 ATGGGAACCAAGTGAAGTGAAGCTCCCTG 842  
RESULT 7  
AF013114 1161 bp mRNA linear ROD 21-AUG-1997  
LOCUS AF013114  
DEFINITION Mus musculus cytokine receptor-like molecule (EBI3) mRNA, complete cds.  
ACCESSION AF013114  
VERSION AF013114.1 GI:2338440  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1161)  
AUTHORS Nomura, H., Yaguchi, N., Hanyuu, C., Maeda, M., Kikuchi, Y., Nakata, Y., Kojima, T., Tulin, E.E. and Hasegawa, M.  
TITLE Mouse homolog of human Epstein-Barr virus-induced gene 3 (EBI3) Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 1161)  
AUTHORS Nomura, H., Yaguchi, N., Hanyuu, C., Maeda, M., Kikuchi, Y., Nakata, Y., Kojima, T., Tulin, E.E. and Hasegawa, M.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUL-1997) Cytokine Research Program, CHUGAI Research Institute for Molecular Medicine, Inc., 153-2, Nagai, Nihari, Ibaraki 300-41, Japan

FEATURES		Location/Qualifiers							
source		1..1161							
		/organism="Mus musculus"							
		/db_xref="taxon:10090"							
		/clone_lib="Clontech Marathon-Ready library #7459-1"							
		/dev_stage="15-day embryo"							
gene		1..1161							
		/gene="EBI3"							
		/note="MMEBI3; MUSEBI3"							
CDS		223..909							
		/gene="EBI3"							
		/note="similar to human EBI3 protein"							
		/codon_start=1							
		/product="cytokine receptor-like molecule"							
		/protein_id="AAB67115.1"							
		/db_xref="GI:2338441"							
		/translation="MSKLFSLALMASRSPGYETETALVALSOPRVQCHASRYPAVD CSWTPLOAPNSTRSTSEIATYRLGVAIQOQSOPCLQSRPQASRCTIPDVHLESTVPM LNTAVHPGASSSLIAFVAERIIKPDPEGVRLRTAGQRLQYLMHPASWPEDIFS LKYRLRYRRRGASHFROVGPFEATFTLRLNSKPHAQCIQVSAQDLTDYGRPSDWSLE GQVESAPHPK"							
	BASE COUNT	244	a	390	c	272	g	255	t
	ORIGIN								
Query Match		40.8%; Score 354; DB 10; Length 1161;							
Best Local Similarity		66.5%; Pred. No. 1.7e-61;							
Matches 539; Conservative		0; Mismatches 265; Indels 6; Gaps 2;							
QY	22	CCCCGAGCTTCTCTGCCCCCTGTCTCTCTGGCCAGCTGCCCGCCCGCAGTGAAGAA	81						
DB	74	CCATCATCATCTCTGCCCCATACACTGGACACTGAGCCACACTGGGAGTCTTCCC	133						
QY	82	AGGGCCCCCAGCAGCTCTGACACTGCCCCGGGTGCAATGCCGAGCCTCTCGGTACCCGAT	141						
DB	134	TGGGGCCAGGTTCCCTGTGTGAGTCCCTGTCTTCAACCTCTCTGTATGGGTCACTAA	193						
QY	142	CGCCGTGATTCCTCTGGACCCCTGCCGATGACCCCGCAGCTTCTCTGGCCCTTGTCC	201						
DB	194	CTCGGATCCAAAGAACAGAGCCACAGAGCATGTCCAAGCTGCTTCTCTGTCACTTGGCC	253						
QY	202	TCTGGGCCAGCTGCCCGCCCTGCGGTGGAAAGAAAGGGCCCCCAGCAGCTCTGACACTGC	261						
DB	254	TCTGGGCCAGCCGCTCCCT---GGTTACACTGAACACAGCTTCTGCTGAGCTTAAGCCAGC	310						
QY	262	CCCGGTGCAATGCCAGCCTCTCGGTACCCGATCGCCGTGATTCCTCTGGACCCCTGC	321						
DB	311	CCAGAGTGCATGCCATGCTTCTCGGTATCCCGTGGCCGTGACTGCTCTGGACTCCTC	370						
QY	322	CGCCTGTCTCCAAACTCCACCAGCCCGCTGTCTTCATTCGCACTGACAGGCTCGCATGG	381						
DB	371	TCCAGGCTCCCAACTCCACAGATCCACGTCCTTCATTCGCACTTACAGGCTCGGTGG	430						
QY	382	CTGCCCCGGGCCACAGCTGGCCTTGCTGACGACGACGCAACGTCACCACTGCACCA	441						
DB	431	CCACCCAGCAGCAGAGCCAGCCCTGCTTACAACAGGAGCCCGCCGATGCACCA	490						
QY	442	TCACGGATGCCAGCTGTTCTTCATGGCTCCCTACGTGCTCAATGTCACCGCCGTCCGC	501						
DB	491	TCCCCGACGTGACCTGTTCTCCACGGTGCCCTACATGCTAAATGTCACGTGACAC	550						
QY	502	CCTGGGGCTCCAGCAGCAGCTTCGTGCTTTCATTAACAGAGACATCATCAAGCCCGACC	561						
DB	551	CAGGCGCGCCAGCAGCAGCCTTCTAGCCTTGTGGTGAAGCAATCATCAAGCCCGACC	610						
QY	562	CTCCAGAGGCGTGCGCCTAAGCCCGCTGCTGAGCGCCAGCTACAGGTGAGTGGAGC	621						
DB	611	CTCCGGAAGGCGTGCGCCTGCGCAGCAGCGGACAGCGC--CTGCAGCTGCTCTGGCATC	667						
QY	622	CTCCCGGGTCTTGCGCCTTCCAGAGATCTTCTCACTGAAGTACTGGATCCGTTACAAGC	681						
DB	668	CCCTGCTTCTTGCGCCTTCCCGGACATCTTCTCTCAAGTACCGACTCCGCTACCGGC	727						
QY	682	GTCAGGAGGCTGCGCGCTTCCACCGGGTGGGCCCATTTGAAGCCACGTCCTTCATCTCA	741						

Db	728	GCCGAGGAGCCTCTCACTTCCGCCAGGTGGACCCATTGAAGCCAGACTTTACCCCTCA	787																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													</
----	-----	--	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

LOCUS AC005578 43297 bp DNA linear PRI 01-SEP-1998  
DEFINITION Homo sapiens chromosome 19, cosmid F20887, complete sequence.  
ACCESSION AC005578  
VERSION AC005578.1 GI:3510237  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 43297)  
AUTHORS Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,  
Burkhardt-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,  
Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Garnes,J.,  
Danganan,L., Poundstone,P., Christensen,M., Georgescu,A., Avila,J.,  
Liu,S., Attix,C., Andreise,T., Trankheim,M., Amico-Keller,G.,  
Coefield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,  
Kronmiller,B., Arellano,A., Montgomery,M., Ow,D., Nolan,M.,  
Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.  
Sequence analysis of a 2.5 Mb region in 19p13.3  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 43297)  
REFERENCE Lamerdin,J.E.  
AUTHORS Direct Submission  
JOURNAL Submitted (01-SEP-1998) Joint Genome Institute, Lawrence Livermore  
COMMENT National Laboratory, 7000 East Ave., Livermore, CA 94551, USA  
Map and sequence oriented from p telomere to centromere. Cosmid  
F20887 overlaps cosmid R33243 to the left by at least 8 kb and  
overlaps BC84616 (CIT-B-144d21) to the right from bases 41,733 to  
43,297. Additional chromosome 19 map and sequence information may  
be obtained at: <http://www-bio.llnl.gov/bbtp/genome/genome.html>.  
FEATURES  
source  
location/Qualifiers  
1. .43297  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19p13.3 between D19S883 and D19S325"  
/clone="F20887"  
/cell\_line="UV5HL9-5B"  
/clone\_lib="LL19NC02 F chromosome 19-specific cosmid  
library"  
/note="Cosmid library constructed at LLNL from flow-sorted  
chromosomes from human-hamster hybrid UV5HL9-5B, which  
carries chromosome 19 as its only human chromosome."  
join(<177. .283,457. .644,1440. .1477,2941. .3048,5246. .5374,  
8654. .8705,8927. .9078,9226. .9936,12267. .12434,  
16803. .16868,17104. .17181)  
/note="Hypothetical human protein most similar to PID  
g1429314 and several ankryn-like proteins"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F20887\_1, partial CDS"  
/protein\_id="AAC33490.1"  
/db\_xref="GI:3510240"  
/translation="SGATPLITAAQMCHTDLCRLLLQGAANDODLOGRTALMLACE  
GASPETVEVLLQGGADPGITDALGQDAHYGALAGDKLILHLQEAQRPSPSALTE  
DDSGEASSODRDAYEEIYRLQERGLQKIRGLEQHKERQEQRLSLHCCSQVQE  
LQQLLVERQEEKESLSGREVESLQSRSLLENRENTSYDVTTLQDEEGPQGLTRPPP  
GAEVLLSRQLSPSAQEHSLAQEQAVAVLTRNQELMEKVQILNFEKDEQOMEVEALA  
EVIPLALYDSLRAEFQDLRQHAELQALRQETREVPREGAAGCESEVAGATATKN  
GPTHMLNGSVAPETKYNGAETIDEAAGDETMARTMEAEATGAETAGLEATGAKVT  
ETKPTGAEVREMETTEEAMMETKPTGAQATDPTTGEVAMGEATTKTAEAEAOAY  
GVGAGQAEPPVTGTTNMEATGSRATGMESTGVSATGVENPGVEATVQREALFMKSERH  
AAEAQLATAEQQLRGLRTEAEARAQOSRAQEAIDKAKEKDKKEAARDHSSVVALYRS  
HLLYAIQGMDEVDVRLISQILQMRLQAQGR"  
repeat\_region  
318. .346  
/rpt\_family="MIR"  
misc\_feature  
457. .644  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 93.000"  
repeat\_region  
complement(1043. .1258)  
/rpt\_family="AluSg"  
1440. .1477  
/note="predicted exon, program: grail2exons\_human\_1.3,

repeat\_region  
complement(1785. .2083)  
/rpt\_family="AluSx"  
repeat\_region  
complement(2124. .2421)  
/rpt\_family="AluSx"  
misc\_feature  
2734. .2814  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 80.000"  
misc\_feature  
2941. .3048  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 96.000"  
repeat\_region  
3114. .3264  
/rpt\_family="LINE2"  
complement(3505. .3808)  
/rpt\_family="AluSg"  
4033. .4330  
/rpt\_family="AluSx"  
repeat\_region  
4495. .4537  
/rpt\_family="LINE2"  
4550. .4587  
/rpt\_family="LINE2"  
repeat\_region  
4595. .4894  
/rpt\_family="AluY"  
4968. .5001  
/rpt\_family="LINE2"  
5246. .5374  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 99.000"  
repeat\_region  
5527. .5817  
/rpt\_family="AluSg"  
complement(5896. .6085)  
/rpt\_family="(GGAA)n"  
repeat\_region  
complement(6086. .6129)  
/rpt\_family="(GGAA)n"  
repeat\_region  
complement(6130. .6323)  
/rpt\_family="AluSg/x"  
repeat\_region  
complement(6326. .6608)  
/rpt\_family="AluSg"  
6700. .6833  
/rpt\_family="(CA)n"  
complement(6845. .7093)  
/rpt\_family="AluJb"  
7249. .7545  
/rpt\_family="AluSx"  
repeat\_region  
7791. .7879  
/rpt\_family="L1"  
7880. .8157  
/rpt\_family="AluSx"  
8161. .8464  
/rpt\_family="AluJb"  
8654. .8705  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 91.000"  
misc\_feature  
8960. .9087  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 100.000"  
9226. .10839  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 92.000"  
complement(10973. .11275)  
/rpt\_family="AluSg"  
complement(11334. .11425)  
/rpt\_family="(GGAA)n"  
complement(11426. .11678)  
/rpt\_family="AluJb"  
11680. .11981  
/rpt\_family="AluSx"  
11982. .12179  
/rpt\_family="AluJb"  
12267. .12434  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 2, quality: excellent, score: 100.000"  
complement(12536. .12569)  
repeat\_region

```

repeat_region /rpt_family="MIR"
complement(12653. .12945)
/rpt_family="AluSx"
repeat_region complement(12957. .13135)
/rpt_family="MIR"
repeat_region complement(13229. .13530)
/rpt_family="AluY"
repeat_region complement(13580. .13751)
/rpt_family="AluJ"
repeat_region complement(13755. .14043)
/rpt_family="AluSx"
repeat_region complement(14074. .14203)
/rpt_family="AluJ"
repeat_region complement(14214. .14318)
/rpt_family="MIR"
repeat_region 14358. .14430
/rpt_family="MIR"
repeat_region complement(14465. .14667)
/rpt_family="MER20"
repeat_region complement(14697. .14731)
/rpt_family="LINE2"
repeat_region 14783. .15083
/rpt_family="AluSx"
repeat_region 15087. .15137
/rpt_family="MER58B"
repeat_region 15178. .15228
/rpt_family="MER58B"
repeat_region complement(15624. .15908)
/rpt_family="AluSg"
repeat_region complement(15919. .16042)
/rpt_family="AluJo/FRAM"
repeat_region 16045. .16077
/rpt_family="(TA)n"
repeat_region complement(16078. .16246)
/rpt_family="AluSx"
repeat_region complement(16251. .16539)
/rpt_family="AluSc"
repeat_region complement(16540. .16675)
/rpt_family="AluSx"
misc_feature 16803. .16868
/note="predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: excellent, score: 89.000"
repeat_region 16896. .16955
/rpt_family="MIR"
misc_feature 17104. .17181
/note="predicted exon, program: grail2exons_human_1.3,
frame: 0, quality: excellent, score: 90.000"
repeat_region 17628. .17925)
/rpt_family="AluY"
repeat_region complement(17926. .18223)
/rpt_family="AluSx"
repeat_region 18327. .18613
/rpt_family="AluJo"
repeat_region complement(18910. .19211)

```

Query Match	21.0%;	Score 182.6;	DB 9;	Length 43297;
Best Local Similarity	89.1%;	Pred. No. 3.7e-27;		
Matches 197; Conservative	0;	Mismatches 24;	Indels 0;	Gaps 0;

QY	334	ACTCCACCAAGCCCCGCTGTCTCTTCATTTGCCACAGCTACAGGCTCGGCATGGCTGCCCCGGGGCC	393
Db	25768	ACTCCCTGAGGCGCCTCAGCGAGGCCCAACCCCTGTGCAGGCTCGGCATGGCTGCCCCGGGGCC	25827
QY	394	ACAGCTGGCCTTGCTTCGACGACGAGCCCAACGTCACACCAGCTGCACCATCAGGATGTCC	453
Db	25828	ACAGCTGGCCCTGCTTCGACGACGAGCCCAACGTCACACCAGCTGCACCATCAGGATGTCC	25887
QY	454	AGCTGTTTCATGCGCTCCCTACGTCGTCATGTTCACCGCGCCGCTCGGGGCTGCA	513
Db	25888	AGCTGTTTCATGCGCTCCCTACGTCGTCATGTTCACCGCGCCGCTCGGGGCTGCA	25947
QY	514	GCAGCAGCTCGTGCCTTTCATTAACAGACATCATCAAG	554

```

Db      25948  GCAGCAGCTTCGTGCTTTTCATACAGAGCACATCAGTGAG  25988

RESULT 10
LOCUS   AC027272/c
DEFINITION
AC027272              179607 bp      DNA          linear      HTG 01-SEP-2000
Homo sapiens chromosome 19 clone RP11-486P13, WORKING DRAFT
SEQUENCE, 18 unordered pieces.
AC027272
AC027272
AC027272.4  GI:9838224
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 179607)
Waterston,R.H.
The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 179607)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 17, 2000 this sequence version replaced gi:8570236.
COMMENT

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0486P13
----- Summary Statistics -----
Sequencing vector: M13; 95%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 95% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170867 bases at least Q40
Consensus quality: 174124 bases at least Q30
Consensus quality: 176036 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 179350; sum-of-contigs
Quality coverage: 4.55 in Q20 bases; agarose-fp
Quality coverage: 4.80 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1338: contig of 1338 bp in length
* 1339 1438: gap of unknown length
* 1439 3330: contig of 1892 bp in length
* 3331 3430: gap of unknown length
* 3431 5429: contig of 1999 bp in length
* 5430 5529: gap of unknown length
* 5530 8107: contig of 2578 bp in length
* 8108 8207: gap of unknown length
* 8208 11180: contig of 2973 bp in length
* 11181 11280: gap of unknown length
* 11281 14039: contig of 2759 bp in length
* 14040 14139: gap of unknown length
* 14140 17230: contig of 3091 bp in length
* 17231 17330: gap of unknown length
* 17331 21334: contig of 4004 bp in length
* 21335 21434: gap of unknown length
* 21435 26877: contig of 5443 bp in length
* 26878 26977: gap of unknown length

```



```
*      26978      31453: contig of 4476 bp in length
*      31454      31553: gap of unknown length
*      31554      36515: contig of 4962 bp in length
*      36516      36615: gap of unknown length
*      36616      46678: contig of 10063 bp in length
*      46679      46778: gap of unknown length
*      46779      57999: contig of 11221 bp in length
*      58000      58099: gap of unknown length
*      58100      74742: contig of 16643 bp in length
*      74743      74842: gap of unknown length
*      74843      92504: contig of 17662 bp in length
*      92505      92604: gap of unknown length
*      92605      119545: contig of 26941 bp in length
*      119546      119645: gap of unknown length
*      119646      147254: contig of 27609 bp in length
*      147255      147354: gap of unknown length
*      147355      179607: contig of 32253 bp in length.

FEATURES
source
1. .179607
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="19"
   /clone="RP11-486P13"
1. .1338
   /note="assembly_name:Contig8"
misc_feature
1439. .3330
   /note="assembly_name:Contig9"
misc_feature
3431. .5429
   /note="assembly_name:Contig10"
misc_feature
5530. .8107
   /note="assembly_name:Contig11"
misc_feature
8208. .11180
   /note="assembly_name:Contig12"
misc_feature
11281. .14039
   /note="assembly_name:Contig13"
misc_feature
14140. .17230
   /note="assembly_name:Contig14"
misc_feature
17331. .21334
   /note="assembly_name:Contig15
clone_end:SP6
vector_side:left"
21435. .26877
   /note="assembly_name:Contig16"
misc_feature
26978. .31453
   /note="assembly_name:Contig17"
misc_feature
31554. .36515
   /note="assembly_name:Contig18"
misc_feature
36616. .46678
   /note="assembly_name:Contig19"
misc_feature
46779. .57999
   /note="assembly_name:Contig20"
misc_feature
58100. .74742
   /note="assembly_name:Contig21"
misc_feature
74843. .92504
   /note="assembly_name:Contig22"
misc_feature
92605. .119545
   /note="assembly_name:Contig23"
misc_feature
119646. .147254
   /note="assembly_name:Contig24"
misc_feature
147355. .179607
   /note="assembly_name:Contig25"
BASE COUNT  41107 a 48085 c 48884 g 39812 t 1719 others
ORIGIN
```

```
Query Match      19.8%; Score 172.2; DB 2; Length 179607;
Best Local Similarity 89.1%; Pred. No. 3.9e-25;
Matches 197; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
```

```
OY      334 ACTCCACGACCCCGTGTCTTCATTGCACGATACAGGCTCGGCATGCTGCCGGGGCC 393
          |||||  |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 158559 ACTCCTGAGCGCTCAGCGAGCCCAACCTGTGCAGGCTCGGCATGCTGCCGGGGCC 158500

OY      394 ACAGCTGGCCTTGCTGCAGCAGACGCCAACGTCACACGCTGCACCATCAGGATGTCC 453
```

```
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 158499 ACAGCTGGCC-TGCCCTGCAGACGAGCGCCAACGTCACACGCTGCACCATCAGGATGTCC 158441

OY      454 AGCTGTTCTCCATGGCTCCCTACGTCGTCATGTCACCGCGCTCCGCCCTGGGGCTCCA 513
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 158440 AGCTGTCTCCATGGCTCCCTACGTCGTCATGTCACCGCGCTCCGCCCTGGGGCTCCA 158381

OY      514 GCAGCAGCTTCGTGCTTTCATACAGACGACATCATCAAG 554
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 158380 GCAGCAGCTTCGTGCTTTCATACAGACGACATCATCAAGTGAAG 158340

RESULT 11
G54090      116 bp      DNA      linear      STS 15-JUL-1999
LOCUS
DEFINITION  902 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION  G54090
VERSION     G54090.1 GI:5499577
KEYWORDS
SOURCE
ORGANISM    Homo sapiens.
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 116)
             Wong,G.K.S., Yu,J., Yang,Z., Hughes,W., Passey,D., Kibukawa,M.,
             Olson,V., Wilson,C. and Green,P.
             Detection and Validation of Single Nucleotide Polymorphisms from
             EST Databases
             Unpublished (1998)

TITLE
JOURNAL
COMMENT
Contact: Gane Ka-Shu Wong
Genome Center
University of Washington
Fluke Hall, Box 352145, Seattle, WA 98195, USA
Tel: 206/685-7348
Fax: 206/685-7344
Email: gksweu.washington.edu
Primer A: CCGAGCCAGTACTACATCC
Primer B: GAGCCCTTGCTACTTGCC
STS size: 116
PCR Profile:
  Presoak: 92 degrees C for 1.00 minute
  Denaturation: 92 degrees C for 0.16 minute
  Annealing: 65 degrees C for 0.50 minute
  Polymerization: 72 degrees C for 1.00 minute
  PCR Cycles: 35
  Thermal Cycler: Perkin Elmer TC
Protocol:
  Template: 7 ng
  Primer: each 0.5 uM
  dNTPs: each 100 uM
  Taq Polymerase: 0.025 units/uL
  Total Vol: 10 uL
Buffer:
  MgCl2: 1.5 mM
  KCl: 50 mM
  Tris-HCl: 10 mM
  pH: 8.3.
Location/Qualifiers
1. .116
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /map="14 CM"
   /clone_lib="Human"
```

```
FEATURES
source
```

```
STS
primer_bind 1. .116
primer_bind 1. .116
BASE COUNT  27 a 36 c 35 g 18 t
ORIGIN
```

```
Query Match      11.8%; Score 102.8; DB 11; Length 116;
Best Local Similarity 98.1%; Pred. No. 8.9e-11;
```

Matches 104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 755 CCGAGCCAGGTACTACGTCCAGTGGCGGCTCAGACCTCACAGACTACGGGGAAGTGA 814  
|||||  
Db 1 CCGAGCCAGGTACTACGTCCAGTGGCGGCTCAGACCTCACAGACTACGGGGAAGTGA 60  
|||||

QY 815 TGACTGAGTCTCCCGCCACTGCCACATGAGCCTGGGCAAGTAG 860  
|||||  
Db 61 TGACTGGAGTCTCCCGCCACTGCCACATGAGCCTGGGCAAGTAG 106  
|||||

RESULT 12  
AC073737/c 193355 bp DNA linear HTG 18-JUL-2000  
LOCUS Mus musculus clone RP23-264G1, WORKING DRAFT SEQUENCE, 24 ordered  
DEFINITION pieces.  
AC073737  
AC073737 GI:9256775  
VERSION HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
KEYWORDS Mus musculus.  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 193355)  
TITLE DOE Joint Genome Institute.  
JOURNAL Sequencing of Mouse  
AUTHORS Unpublished  
TITLE 2 (bases 1 to 193355)  
REFERENCE DOE Joint Genome Institute.  
AUTHORS Direct Submission  
TITLE Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 18, 2000 this sequence version replaced gi:8810354.  
COMMENT  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1838393  
Center clone name: RPCI-23\_264G1  
-----

Summary Statistics  
Consensus quality: 175100 bases at least Q40  
Consensus quality: 187679 bases at least Q30  
Consensus quality: 189929 bases at least Q20  
Estimated insert size: 87000; agarose-fp estimation  
Estimated insert size: 192255; sum-of-contrigs estimation  
Quality coverage: 10.37 in Q20 bases; agarose-fp estimation  
Quality coverage: 4.69 in Q20 bases; sum-of-contrigs estimation  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 24 contrigs. Gaps between the contrigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 10615: contig of 10615 bp in length  
10616 10715: gap of unknown length  
10716 40728: contig of 30013 bp in length  
40729 40828: gap of unknown length  
40829 43385: contig of 2557 bp in length  
43386 43485: gap of unknown length  
43486 60463: contig of 16978 bp in length  
60464 60563: gap of unknown length  
60564 63757: contig of 3194 bp in length  
63758 63857: gap of unknown length  
63858 72484: contig of 8627 bp in length  
72485 72584: gap of unknown length  
72585 80251: contig of 7667 bp in length  
80252 80351: gap of unknown length

80352 101558: contig of 21207 bp in length  
\* 101559 101658: gap of unknown length  
\* 101659 113941: contig of 12283 bp in length  
\* 113942 114041: gap of unknown length  
\* 114042 118302: contig of 4261 bp in length  
\* 118303 118402: gap of unknown length  
\* 118403 122141: contig of 3739 bp in length  
\* 122142 122241: gap of unknown length  
\* 122242 141898: contig of 19657 bp in length  
\* 141899 141998: gap of unknown length  
\* 141999 145605: contig of 3607 bp in length  
\* 145606 145705: gap of unknown length  
\* 145706 148555: contig of 2850 bp in length  
\* 148556 148655: gap of unknown length  
\* 148656 151425: contig of 2770 bp in length  
\* 151426 151525: gap of unknown length  
\* 151526 158286: contig of 6761 bp in length  
\* 158287 158386: gap of unknown length  
\* 158387 160638: contig of 2252 bp in length  
\* 160639 160738: gap of unknown length  
\* 160739 173388: contig of 12650 bp in length  
\* 173389 173488: gap of unknown length  
\* 173489 179077: contig of 5589 bp in length  
\* 179078 179177: gap of unknown length  
\* 179178 184518: contig of 5341 bp in length  
\* 184519 184618: gap of unknown length  
\* 184619 189175: contig of 4557 bp in length  
\* 189176 189275: gap of unknown length  
\* 189276 191154: contig of 1879 bp in length  
\* 191155 191254: gap of unknown length  
\* 191255 192103: contig of 849 bp in length  
\* 192104 192203: gap of unknown length  
\* 192204 193355: contig of 1152 bp in length.  
FEATURES  
source location/Qualifiers  
1..193355  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-264G1"  
/clone\_lib="RPCI mouse BAC library 23"  
BASE COUNT 49405 a 44453 c 45086 g 52104 t 2307 others  
ORIGIN  
Query Match 11.2%; Score 97; DB 2; Length 193355;  
Best Local Similarity 68.9%; Pred. No. 5.4e-10;  
Matches 133; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
QY 362 CACGTACAGGCTCGGCATGGCTGCCCGGGCCACAGCTGGCGCTTGCTGCAGACAGCGCC 421  
Db 111878 CCCATGCAGGCTCGGTGTGGCCACCACAGCAGACAGAGCCCTGCTTACACAGAGCC 111819  
QY 422 AACGTCACAGCTGCACCATCACGGATGTCAGCTGTCTTCATAGCTCCCTACGTGCT 481  
Db 111818 CCAGGCTCCCGATGCACCATCCCGCAGCGTGCACCTGTCTTCACGGTGCCCTACATGCT 111759  
QY 482 CAATGTCACCGCGCTCCGCCCTGGGGCTCCAGCAGCAGCTTCGTCCTTTCATACAGA 541  
Db 111758 AAATGTCAGTGCAGTGCACCCAGCGGGCGCCAGCAGCAGCCTCTTGTGGCTGA 111699  
QY 542 GCACATCATCAAG 554  
Db 111698 GCGAATCAGTGAG 111686  
RESULT 13  
AC118074/c 176359 bp DNA linear HTG 18-JUL-2002  
LOCUS Rattus norvegicus clone CH230-47J17, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 69 unordered pieces.  
AC118074  
AC118074 GI:21746302  
VERSION HTG; HTGS\_PHASE1.  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 176359)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unpublished  
2 (bases 1 to 176359)  
Worley,K.C.  
Direct Submission  
Submitted (13-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 176359)  
Worley,K.C.  
Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20177536.  
COMMENT  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center Project name: GUEB  
Center Clone name: CH230-47J17  
----- Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 115372 bases at least Q40  
Consensus quality: 121114 bases at least Q30

Consensus quality: 124422 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 69 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
1220 1219: contig of 1219 bp in length  
\* 1320 2744: contig of 1425 bp in length  
\* 2745 2844: gap of unknown length  
\* 2845 4066: contig of 1222 bp in length  
\* 4067 4166: gap of unknown length  
\* 4167 5794: contig of 1628 bp in length  
\* 5795 5894: gap of unknown length  
\* 5895 6936: contig of 1042 bp in length  
\* 6937 7036: contig of 1042 bp in length  
\* 7037 8767: gap of unknown length  
\* 8768 8867: contig of 1731 bp in length  
\* 8868 10546: gap of unknown length  
\* 10547 10646: contig of 1679 bp in length  
\* 10647 12271: gap of unknown length  
\* 12272 12371: contig of 1625 bp in length  
\* 12372 13483: gap of unknown length  
\* 13483 13583: contig of 1112 bp in length  
\* 13584 14951: gap of unknown length  
\* 13584 14951: gap of unknown length  
\* 14952 15051: contig of 1368 bp in length  
\* 14952 15051: gap of unknown length  
\* 15052 16270: contig of 1219 bp in length  
\* 15052 16270: contig of 1219 bp in length  
\* 16271 16370: gap of unknown length  
\* 16371 17804: contig of 1434 bp in length  
\* 17805 17904: gap of unknown length  
\* 17905 19394: contig of 1490 bp in length  
\* 19395 19494: gap of unknown length  
\* 19495 20712: gap of unknown length  
\* 20713 20812: contig of 1218 bp in length  
\* 20813 22799: gap of unknown length  
\* 22800 22899: contig of 1987 bp in length  
\* 22900 24177: gap of unknown length  
\* 24178 24277: gap of unknown length  
\* 24278 25821: contig of 1278 bp in length  
\* 25821 25921: gap of unknown length  
\* 25922 27784: contig of 1544 bp in length  
\* 27785 27884: gap of unknown length  
\* 27885 28934: contig of 1863 bp in length  
\* 28935 29034: gap of unknown length  
\* 29035 30354: contig of 1050 bp in length  
\* 30355 30454: gap of unknown length  
\* 30455 31701: contig of 1320 bp in length  
\* 31702 31801: gap of unknown length  
\* 31802 33337: contig of 1247 bp in length  
\* 33338 33437: gap of unknown length  
\* 33438 34728: contig of 1536 bp in length  
\* 34729 34828: gap of unknown length  
\* 34829 36591: contig of 1291 bp in length  
\* 36592 36691: gap of unknown length  
\* 36692 38003: contig of 1763 bp in length  
\* 38004 38103: gap of unknown length  
\* 38104 40127: contig of 1312 bp in length  
\* 40128 40227: gap of unknown length  
\* 40228 41691: contig of 2024 bp in length  
\* 41692 41791: gap of unknown length  
\* 41792 43535: contig of 1464 bp in length  
\* 43536 43635: gap of unknown length  
\* 43636 45466: contig of 1744 bp in length  
\* 45467 45566: gap of unknown length  
\* 45567 47021: contig of 1831 bp in length  
\* 47022 47121: gap of unknown length  
\* 47122 48801: contig of 1455 bp in length  
\* 48801: contig of 1680 bp in length

*	48802	48901: gap of unknown length
*	48902	49935: contig of 1034 bp in length
*	49936	50035: gap of unknown length
*	50036	51484: contig of 1449 bp in length
*	51485	51584: gap of unknown length
*	51585	53366: contig of 1782 bp in length
*	53367	53466: gap of unknown length
*	53467	55706: contig of 2240 bp in length
*	55707	55806: gap of unknown length
*	55807	57484: contig of 1678 bp in length
*	57485	57584: gap of unknown length
*	57585	59524: contig of 1940 bp in length
*	59525	59624: gap of unknown length
*	59625	61504: contig of 1880 bp in length
*	61505	61604: gap of unknown length
*	61605	63038: contig of 1434 bp in length
*	63039	63138: gap of unknown length
*	63139	65148: contig of 2010 bp in length
*	65149	65248: gap of unknown length
*	65249	68165: contig of 2917 bp in length
*	68166	68265: gap of unknown length
*	68266	70437: contig of 2172 bp in length
*	70438	70537: gap of unknown length
*	70538	72778: contig of 2241 bp in length
*	72779	72878: gap of unknown length
*	72879	74865: contig of 1987 bp in length
*	74866	74965: gap of unknown length
*	74966	78152: contig of 3187 bp in length
*	78153	78252: gap of unknown length
*	78253	80830: contig of 2578 bp in length
*	80831	80930: gap of unknown length
*	80931	83775: contig of 2845 bp in length
*	83776	83875: gap of unknown length
*	83876	86764: contig of 2889 bp in length
*	86765	86864: gap of unknown length
*	86865	90526: contig of 3662 bp in length
*	90527	90626: gap of unknown length
*	90627	92648: contig of 2022 bp in length
*	92649	92748: gap of unknown length
*	92749	96414: contig of 3666 bp in length
*	96415	96514: gap of unknown length
*	96515	99630: contig of 3116 bp in length
*	99631	99730: gap of unknown length
*	99731	102795: contig of 3065 bp in length
*	102796	102895: gap of unknown length
*	102896	105478: contig of 2583 bp in length

Query Match 10.4%; Score 90.2; DB 2; Length 176359;  
Best Local Similarity 67.0%; Pred. No. 1.3e-08;  
Matches 128; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY	364	CGTACAGGCTCGGCATGGCTGCCCGGGGCGCACAGCTGGCCTTGCTGCAGCAGACGCCAA	423
Db	66695	CATGCAGGCTCGGTGTGGCCGCCAGCAGACGACGCCCTGCTGCAACGGAACCCCCC	66636
QY	424	CGTCCACGAGCTGCACCATCACGGATGTCCAGCTGTCTCCATGGCTCCCTAGTGCTCA	483
Db	66635	AGGCATCCCGTTGCACCATCCGCCACGTCACCTGTTTCCACAGTGGCCCTACTTGCTAA	66576
QY	484	ATGTACCGCGCTCGGCCCTGGGGCTCCAGCAGCAGCTTGCTTCATACAGAGC	543
Db	66575	ATGTACATGTAGTGACACCTGGCGCGCAACAGCAGACCTTCTAGCCTTTGTGGCGGAAC	66516
QY	544	ACATCAATCAAG	554
Db	66515	GAATCAGTAAG	66505

RESULT 14  
LOCUS AX399177  
DEFINITION Sequence 17 from Patent WO0194605.  
ACCESSION AX399177

1332 bp DNA linear PAT 27-MAY-2002

VERSION AX399177.1 GI:21261511  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT 291 a 428 c 338 g 275 t  
ORIGIN

1  
Zolotukhin,S., Kalra,S.P., Moldawer,L.L., Muzyczka,N.,  
Hauswirth,W.W. and Kalra,P.S.  
Aav vector-based compositions and therapeutic uses thereof  
Patent: WO 0194605-A 17 13-DEC-2001;  
University of Florida (US)  
Location/Qualifiers  
1. 1332  
/organism="Rattus sp."  
/db\_xref="taxon:10118"

Query Match 8.6%; Score 74.8; DB 6; Length 1332;  
Best Local Similarity 52.1%; Pred. No. 2.9e-05;  
Matches 219; Conservative 0; Mismatches 192; Indels 9; Gaps 2;

QY	429	ACCAGCTGCACCATACGGATGTCCAGCTGTCTCCATGGCTCCCTACGTGCTCATGTG	488
Db	564	AACCGCTGTCAATCGATCGTACATGACACCTGTCTCAACCATCAAGTACAAAGTCTCCATA	623
QY	489	ACCGCCGTCGCGCCCGGGGCTCCAGCAGCAGCTTGTGCTTTCATTAACAGACATC	548
Db	624	AGTGTACGACACGCTTGGGTACACAACACACGCGCTATCACCTTGAGCAATTCACCATT	683
QY	549	ATCAAGCCCGACCCCTCCAGAAGCGTG-----CGCCTAAGCCCCCTGCTGAGCGCCAG	602
Db	684	GTCAGCCCGATCCTCCAGAAATGTGTGGCCCGGCCAGTGCCAGCAACCCCGTCA	743
QY	603	CTACAGGTGCAGTGGAGGCTCCCGGCTGCTGCGCCTTCCAGAGATCTTCACTGAAG	662
Db	744	CTGGAGGTGACATGGCAGACCCCTCAACTTGGCTGATCCGAACTTTCACCTCAAG	803
QY	663	TACTGCATCCGTTTACAGCGTCAGGAGCTGCGGCTTCCACCGGCTGGGCCATTGAA	722
Db	804	TTTTTTCTGCGCTACCGGCTCTCATCTGGATCAATGGCAGCATGTGAGCTCTCGAAT	863
QY	723	GCCACGTCCTTCATCTCTCAGGCGCTGCGGCCCGCCAGCCAGGTACTACGTCAAGTGCGG	782
Db	864	GGCACAGCCCAACACCATCACGGATGCTATGCTGGGAGGAGATCATCATCAGGTGGCC	923
QY	783	GCTCAGACCTTCACAGACTACGGGGAAGTGAAGTGAAGTCTCCCGCCACTGCCACA	842
Db	924	GCCAAGGA---CAATGAGATTGGAGCATGAGTGAAGTGAAGTGTGGCTGCACGCCACA	980

RESULT 15  
LOCUS S54212  
DEFINITION 1332 bp mRNA linear ROD 08-MAY-1993  
ACCESSION S54212  
VERSION S54212  
KEYWORDS S54212  
SOURCE S54212.1 GI:265055  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1  
S54212  
ciliary neurotrophic factor receptor alpha component [rats, brain,  
mRNA, 1332 nt].  
S54212  
S54212.1 GI:265055  
Rattus sp. brain.  
Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1332)  
Ip,N.Y., McClain,J., Barrezueta,N.X., Aldrich,T.H., Pan,L., Li,Y.,  
Wiegand,S.J., Friedman,B., Davis,S. and Yancopoulos,G.D.  
The alpha component of the CNTF receptor is required for signaling  
and defines potential CNTF targets in the adult and during  
development  
Neuron 10 (1), 89-102 (1993)





•  
•  
•  
•

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 28, 2002, 10:16:24 ; Search time 1716.5 Seconds

(without alignments)  
3255.309 Million cell updates/sec

Title: US-09-921-994-2  
Perfect score: 1064  
Sequence: 1 MTPQLLLALVLMASCPPRSG.....HNRAHHQARPSRRRAPKPPR 192

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09921994/runat\_27122002\_084132\_6883/app\_query.fasta\_1.782  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09921994.@CGN\_1\_1\_3637\_@runat\_27122002\_084132\_6883 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADES=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1064	100.0	868	6	AX377958	AX377958 Sequence
2	784.5	73.7	1195	9	BC015364	BC015364 Homo sapi
3	782	73.5	1161	6	AR003351	AR003351 Sequence
4	782	73.5	1161	6	AR050784	AR050784 Sequence
5	782	73.5	1161	9	HUMEB13X	L08187 Human cyto
6	374.5	35.2	43297	9	AC005578	AC005578 Homo sapi
7	373	35.1	445	6	AX392987	AX392987 Sequence
8	351.5	33.0	179607	2	AC027272	AC027272 Homo sapi
9	281.5	26.5	1131	10	BC008209	BC008209 Mus muscu
10	281.5	26.5	1161	10	AF013114	AF013114 Mus muscu
11	193.5	18.2	193355	2	AC073737	AC073737 Mus muscu
12	173.5	16.3	176359	2	AC118074	AC118074 Rattus no
13	149.5	14.1	10770	2	AE012121	AE012121 Xanthomon
14	149	14.0	169964	2	AC131212	AC131212 Homo sapi
15	149	14.0	186558	2	AC079031	AC079031 Homo sapi
16	148	13.9	1969	6	I93634	I93634 Sequence 3
17	146.5	13.8	82746	1	AF453501	AF453501 Actinosyn
18	146	13.7	133330	2	OSJN00266	AL731624 Oryza sat
19	145.5	13.7	22449	1	SPSNBCDE	Y11548 S.pristinae
20	145.5	13.7	22449	1	SPSNBCGEN	X98690 S.pristinae
21	145	13.6	1850	1	ACVFTMBA	M32067 A.viscosus
22	145	13.6	1969	6	I93632	I93632 Sequence 1
23	145	13.6	1969	6	I93633	I93633 Sequence 2
24	145	13.6	2664	8	ZMAHAS108	X63553 Z.mays gene
25	144	13.5	38390	1	SC2H12	AL359215 Streptomy
26	143.5	13.5	66618	9	HSU579N16	AL096767 Human DNA
27	143	13.4	132914	2	CNS08CAA	AL831797 Oryza sat
28	142.5	13.4	125505	2	AP004177	AP004177 Oryza sat
29	142	13.3	187757	2	AC087557	AC087557 Mus muscu
30	141.5	13.3	169230	9	AP002505	AP002505 Homo sapi
31	141.5	13.3	188604	9	AP002439	AP002439 Homo sapi
32	141	13.3	14472	1	AE003928	AE003928 Xylella f
33	141	13.3	18959	1	AE004032	AE004032 Xylella f
34	141	13.3	20956	1	AE004082	AE004082 Xylella f
35	140.5	13.2	135789	8	AC051634	AC051634 Oryza sat
36	140	13.2	4257	6	AR144708	AR144708 Sequence
37	140	13.2	4257	6	AR145616	AR145616 Sequence
38	140	13.2	6633	14	HEHSVIG3	X06461 Herpes simp
39	140	13.2	12001	6	AR048721	AR048721 Sequence
40	140	13.2	26245	14	HS1US	L00036 Human herpe
41	140	13.2	26245	14	HS1US	L00036 Human herpe
42	140	13.2	38425	1	SCPF55	AL132991 Streptomy
43	140	13.2	152261	14	HE1CG	X14112 Herpes simp
44	140	13.2	152261	14	HE1CG	X14112 Herpes simp
45	139.5	13.1	1466	5	AF091043	AF091043 Coturnix

RESULT 1

ALIGNMENTS

```
AX377958      AX377958      868 bp      DNA      linear      PAT 18-MAR-2002
LOCUS         AX377958      Sequence 1 from Patent WO0212282.
DEFINITION    AX377958
ACCESSION     AX377958
VERSION       AX377958.1  GI:19574012
KEYWORDS      '
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS       Bowman,M.R.
TITLE         Novel ebi-3-alt protein and nucleic acid molecules and uses
               therefor
JOURNAL       Patent: WO 0212282-A 1 14-FEB-2002;
GENETICS      INSTITUTE, INC. (US)
FEATURES      location/Qualifiers
               source          1..868
                               /organism="Homo sapiens"
                               /db_xref="taxon:9606"
BASE COUNT    147 a      325 c      239 g      157 t
ORIGIN
Alignment Scores:
Pred. No.:    6.79e-34      Length:      868
Score:         1064.00      Matches:    192
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:   100.00%      Indels:     0
DB:            Gaps:        0

US-09-921-994-2 (1-192) x AX377958 (1-868)
QY      1 MetThrProGlnLeuLeuLeuAlaLeuValLeuTrpAlaSerCysProProArgSerGly 20
        |||||||
Db      17 ATGACCCCGCAGCTTCTCTGCGCCCTTGCTCTCTGGGCCAGCTGCCGCCGCCGACATGGA 76
QY      21 ArgLysGlyProProAlaAlaLeuThrLeuProArgValGlnCysArgAlaSerArgTyr 40
        |||||||
Db      77 AGGAAAGGGCCCCAGCAGCTCTGACACTGCCCGGGTGCAATGCCGAGCCTCTCGGTAC 136
QY      41 ProIleAlaValAspCysSerTrpThrLeuProHisAspProAlaAlaSerProGlyPro 60
        |||||||
Db      137 CCGATCGCCGTGATGCTCTCTGACCCCTGCGCATGACCCCGCAGCTTCTCTGCGCCCT 196
QY      61 CysProLeuGlyGlnLeuProAlaLeuArgTrpLysGluArgAlaProSerSerSerAsp 80
        |||||||
Db      197 TGTCTCTGTGGCCAGCTGCCCCCGCTGCGGTGAAGAAAGGGCCCCCAGCAGCTCTGAC 256
QY      81 ThrAlaProGlyAlaMetProSerLeuSerValProAspArgArgGlyLeuLeuLeuAsp 100
        |||||||
Db      257 ACTGCCCCGGGTGCAATGCCGAGACCTCTCGGTACCCGATCGCCGTGATTCCTCTGAC 316
QY      101 ProAlaAlaCysSerLysLeuHisGlnProArgValLeuHisCysHisValGlnAlaArg 120
        |||||||
Db      317 CCTGCCCGCTGCTCCAACCTCCACACGCCCCGTCCTTCATTGGCCACGTACAGGCTCGG 376
QY      121 HisGlyCysProGlyProGlnLeuAlaLeuProAlaAlaAspAlaAsnValHisGlnLeu 140
        |||||||
Db      377 CATGGCTGCCCGGGGCCACAGCTGGCCTTGCCTCGACAGACGCCAACGTCCACACAGCTG 436
QY      141 HisHisHisGlyCysProAlaValLeuHisGlySerLeuArgAlaGlnCysHisArgArg 160
        |||||||
Db      437 CACCATCAGGATGTCCAGCTGTCTCATGTGCTCCCTACGTGCTCAATGTCAACGCCCGT 496
QY      161 ProProLeuGlyLeuGlnGlnGlnLeuArgAlaPheHisAsnArgAlaHisHisGlnAla 180
        |||||||
Db      497 CCGCCCCCTGGGGCTCCAGCAGCAGCTTCGTGCTTCATTACAAGACACATCATCAGCC 556
QY      181 ArgProSerArgArgArgAlaProLysProProArg 192
        |||||||
Db      557 CGACCCCTCAGAAGCGGTGCGCCTAAGCCCCCTCGC 592
```

```
RESULT 2
BC015364
LOCUS         BC015364      1195 bp      mRNA      linear      PRI 04-OCT-2001
DEFINITION    Homo sapiens, Epstein-Barr virus induced gene 3, clone MGC:21515
IMAGE:3897136, mRNA, complete cds.
ACCESSION     BC015364
VERSION       BC015364.1  GI:15929884
KEYWORDS      MGC.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS       Strausberg,R.
TITLE         Direct Submission
JOURNAL       Submitted (01-OCT-2001) National Institutes of Health, Mammalian
               Gene Collection (MGC), Cancer Genomics Office, National Cancer
               Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
               USA
REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT        Contact: MGC help desk
               Email: cgabs-remail.nih.gov
               Tissue Procurement: ATCC
               CDNA Library Preparation: Life Technologies, Inc.
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Sequencing Group at the Stanford Human
               Center, Stanford University School of Medicine, Stanford, CA 94305
               Web site:      http://www-shgc.stanford.edu
               Contact: (Dickson, Mark) mcd@paxil.stanford.edu
               Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
               R. M.
               Clone distribution: MGC clone distribution information can be found
               through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
               Series: IRAX Plate: 22 Row: f Column: 18
               This clone was selected for full length sequencing because it
               passed the following selection criteria: Similarity but not
               identity to protein.
FEATURES      location/Qualifiers
               source          1..1195
                               /organism="Homo sapiens"
                               /db_xref="taxon:9606"
                               /db_xref="taxon:9606"
                               /clone="MGC:21515 IMAGE:3897136"
                               /tissue_type="Pancreas, epithelioid carcinoma"
                               /clone_1lb="NIH_MGC_70"
                               /lab_host="DH10B"
                               /note="Vector: pCMV-SPORT6"
                               54..743
                               /codon_start=1
                               /product="Epstein-Barr virus induced gene 3"
                               /protein_id="AAH15364.1"
                               /db_xref="GI:15929885"
                               /translation="MTPQLLLALVLMASCPSCGRKGPAPALTLPVQCRASRYPIAV
                               DCSWTLPPAPNSTSPVSFIATYRLGMAARGHSWPCLOQTPTSTCTITDVQLFSMAPI
                               VLNVTAVHPWSSSSSFVPTTEHIKRPDEGVRLSPLAERQVQWPEPGSWPPEI
                               FSLKWIRYKRQGAARFHRVGPTEATSFILRAVPRARYVQVAAQDLTDYGELSDWS
                               LPATATMSIGK"
BASE COUNT    247 a      395 c      311 g      242 t
ORIGIN
Alignment Scores:
Pred. No.:    5.94e-23      Length:      1195
Score:         784.50      Matches:    144
Percent Similarity: 90.12%  Conservative: 2
Best Local Similarity: 88.89%  Mismatches: 11
Query Match:   73.73%      Indels:     1
DB:            Gaps:        1

US-09-921-994-2 (1-192) x BC015364 (1-1195)
QY      31 ProArgValGlnCysArgAlaSerArgTyrProIleAlaValAspCysSerTrpThrLeu 50
        |||
```



Db 23 CCCACTCCTGAGAGCAGAGCTGGCCGC-----49

QY 51 ProHisAspProAlaAlaSerProGlyProCysProLeuGlyGlnLeuProAlaLeuArg 70  
|||||  
Db 50 AGCCATGACCCCGCAGCTTCTCTGCGCCCTTGCTCTGTGGCCAGCTGCCCGCCCTGAG 109

QY 71 TrpLysGluArgAlaProSerSerSerAspThrAlaProGlyAlaMetProSerLeuSer 90  
|||||  
Db 110 TGGAAAGAAAGGGCCCCCAGAGCTGTGACACTGGCCCCGGGTGCAATGCCGAGACTCTCG 169

QY 91 ValProAspArgArgGlyLeuLeuLeuAspProAlaAlaCysSerLysLeuHisGlnPro 110  
|||||  
Db 170 GTACCCGATCGCCGCTGATGCTGTGGACCCCTGCGCCCTGCTCCAACCTCCACACGCC 229

QY 111 ArgValLeuHisCysHisValGlnAlaArgHisGlyCysProGlyProGlnLeuAlaLeu 130  
|||||  
Db 230 CGTGCTCTTCATTGCCACGTACAGGCTCGGCATGGCTGCCGGGGCCACAGCTGGCCCTG 289

QY 131 ProAlaAlaAspAlaAsnValHisGlnLeuHisHisGlyCysProAlaValLeuHis 150  
|||||  
Db 290 CCTGCAGCAGACGCCCAACGTCACACAGCTGCACCATCATCGGATGTCCAGCTGTCTCCAT 349

QY 151 GlySerLeuArgAlaGlnCysHisArgArgProProLeuGlyLeuGlnGlnLeuArg 170  
|||||  
Db 350 GGCCTCCCTACGTGCTCAATGTCAACGCCGCTCCACCCCTGGGGCTCCAGCAGCAGCTTCGT 409

QY 171 AlaPheHisAsnArgAlaHisHisGlnAlaArgProSerArgArgAlaProLysPro 190  
|||||  
Db 410 GCCTTTCATACAGACAGACATCATCAAGCCGACCCCTCCAGAAAGCGGTGCGCTTAAGCCC 469

QY 191 ProArg 192  
|||||

Db 470 CCTCGC 475

RESULT 3  
AR003351

LOCUS AR003351 1161 bp DNA linear PAT 04-DEC-1998

DEFINITION Sequence 5 from patent US 5744301.

ACCESSION AR003351

VERSION AR003351.1 GI:3964610

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1161)  
AUTHORS Birkenbach,M. and Kieff,E.  
TITLE Methods of detection of Epstein barr virus induced genes expressed in the placenta  
JOURNAL Patent: US 5744301-A 5 28-APR-1998;  
FEATURES Location/Qualifiers  
source 1..1161

BASE COUNT 239 a 378 c 304 g 240 t

ORIGIN

Alignment Scores:  
Pred. No.: 7.24e-23 Length: 1161  
Score: 782.00 Matches: 140  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.29% Mismatches: 0  
Query Match: 73.50% Indels: 0  
DB: 6 Gaps: 0

US-09-921-994-2 (1-192) x AR003351 (1-1161)

QY 52 HisAspProAlaAlaSerProGlyProCysProLeuGlyGlnLeuProAlaLeuArgTrp 71  
|||||  
Db 13 CATGACCCCGCAGCTTCTCTGCGCCCTGTCTCTGGGCCAGCTGCCCGCCCTGCAGTGG 72

QY 72 LysGluArgAlaProSerSerSerAspThrAlaProGlyAlaMetProSerLeuSerVal 91  
|||||  
Db 73 AAGGAAGGGCCCCCAGCAGCTGTGACACTGCCCCGGGTGCAATGCCGAGCCTCTCGGTA 132

QY 92 ProAspArgGlyLeuLeuLeuAspProAlaAlaCysSerLysLeuHisGlnProArg 111  
|||||  
Db 133 CCCGATCGCCGCTGGATTGCTCTCGTGGACCCCTGCCCTGTCCAAACTCCACAGCCCCGT 192

QY 112 ValLeuHisCysHisValGlnAlaArgHisGlyCysProGlyProGlnLeuAlaLeuPro 131  
|||||  
Db 193 GTCCTTCATTGCCACGTACAGGCTCGGCATGGCTGCCGGGGCCACAGCTGGCCCTGCT 252

QY 132 AlaAlaAspAlaAsnValHisGlnLeuHisHisGlyCysProAlaValLeuHisGly 151  
|||||  
Db 253 GCAGCAGACGCCCAACGTCACACAGCTGCACCATCAAGGATGTCCAGCTGTCTCCATGGC 312

QY 152 SerLeuArgAlaGlnCysHisArgArgProProLeuGlyLeuGlnGlnLeuArgAla 171  
|||||  
Db 313 TCCCTACGTGCTCAATGTCAACGCCGCTCCACCCCTGGGGCTCCAGCAGCAGCTTCGTGCC 372

QY 172 PheHisAsnArgAlaHisHisGlnAlaArgProSerArgArgAlaProLysProPro 191  
|||||  
Db 373 TTTCATTAACAGACAGACATCATCAAGCCCGAACCCCTCCAGAAAGCGGTGCGCCTAAGCCCCCT 432

QY 192 Arg 192  
||||

Db 433 CGC 435

RESULT 4  
AR050784

LOCUS AR050784 1161 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5830451.

ACCESSION AR050784

VERSION AR050784.1 GI:5974148

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1161)  
AUTHORS Devergne,O. and Kieff,E.D.  
TITLE Haematopoietic cytokine Epstein Barr virus-induced protein  
JOURNAL Patent: US 5830451-A 1 03-NOV-1998;  
FEATURES Location/Qualifiers  
source 1..1161

BASE COUNT 239 a 378 c 304 g 240 t

ORIGIN

Alignment Scores:  
Pred. No.: 7.24e-23 Length: 1161  
Score: 782.00 Matches: 140  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.29% Mismatches: 0  
Query Match: 73.50% Indels: 0  
DB: 6 Gaps: 0

US-09-921-994-2 (1-192) x AR050784 (1-1161)

QY 52 HisAspProAlaAlaSerProGlyProCysProLeuGlyGlnLeuProAlaLeuArgTrp 71  
|||||  
Db 13 CATGACCCCGCAGCTTCTCTGCGCCCTGTCTCTGGGCCAGCTGCCCGCCCTGCAGTGG 72

QY 72 LysGluArgAlaProSerSerSerAspThrAlaProGlyAlaMetProSerLeuSerVal 91  
|||||  
Db 73 AAGGAAGGGCCCCCAGCAGCTGTGACACTGCCCGGGGTGCAATGCCGAGCCTCTCGGTA 132

QY 92 ProAspArgArgGlyLeuLeuLeuAspProAlaAlaCysSerLysLeuHisGlnProArg 111  
|||||  
Db 133 CCCGATCGCCGCTGGATTGCTCTGGACCCCTGCCCTGTCCAAACTCCACACGCCCCGT 192

QY 112 ValLeuHisCysHisValGlnAlaArgHisGlyCysProGlyProGlnLeuAlaLeuPro 131  
|||||  
Db 193 GTCCTTCATTGCCACGTACAGGCTCGGCATGGCTGCCGGGGCCACAGCTGGCCCTGCT 252

QY 132 AlaAlaAspAlaAsnValHisGlnLeuHisHisGlyCysProAlaValLeuHisGly 151  
|||||  
Db 253 GCAGCAGACGCCCAACGTCACACAGCTGCACCATCAAGGATGTCCAGCTGTCTCCATGGC 312

OY	152	SerLeuArgAlaGlnCysHisArgArgProProLeuGlyLeuGlnGlnInLeuArgAla	171
DB	313	TCCCTACGTGCTCAATGTCACCGCCGTCACACCCTGGGGCTCCAGCAGACTTGC GTGC	372
OY	172	PheHisAsnArgAlaHisHisGlnAlaArgProSerArgArgArgAlaProLysProPro	191
DB	373	TTTCATAACAGACGACATCATCAAGCCCGACCTCCAGAAGCGTGCGCTTAAGCCCCCT	432
OY	192	Arg	192
DB	433	CGC	435
RESULT	5		
LOCUS	HUMEBI3X	1161 bp	mRNA linear PRI 01-APR-1996
DEFINITION	Human cytokine receptor (EBI3) mRNA, complete cds.		
ACCESSION	L08187		
VERSION	L08187.1	GI:632973	
KEYWORDS	Cytokine receptor.		
SOURCE	Homo sapiens cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Devergne,O., Hummel,M., Koepfen,H., Le Beau,M.M., Nathanson,E.C., Kieff,E. and Birkenbach,M.		
TITLE	A novel interleukin-12 p40-related protein induced by latent Epstein-Barr virus infection in B lymphocytes		
JOURNAL	J. Virol.	70 (2), 1143-1153	(1996)
MEDLINE	96135230		
PUBMED	8551575		
COMMENT	On Jan 24, 1995 this sequence version replaced gi:181921.		
FEATURES	Location/Qualifiers		
source	1. .1161		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/cell_line="BL41/B95-8"		
	/cell_type="B lymphocyte, EBV-converted Burkitt lymphoma"		
gene	1. .1161		
	/gene="EBI3"		
CDS	14. .703		
	/gene="EBI3"		
	/codon_start=1		
	/product="cytokine receptor"		
	/protein_id="AAA93193.1"		
	/db_xref="GI:632974"		
	/translation="MTPQLLALVLWASCPPCSGRKGPAAALTLPVQCRASRYPIAVDCSWTLPPAPNSTSPVSFIATYRLGMARGHSWCLQOTPTSTCTITDVQLFSMAPYVLNVTAVHPWGSSSFVPFITEHTIKPDPEGVRLSPLAERHVQVQWEPPGSWFPEIFSLKWIIRYKRKGGAARFHRVGPIEATSFILRAVRPRARYVVQAQDLTDYGELSDWSLPATATMSLGK"		
	14. .73		
sig_peptide	/gene="EBI3"		
mat_peptide	74. .700		
	/gene="EBI3"		
	/product="cytokine receptor"		
BASE COUNT	239 a 378 c 304 g 240 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	7.24e-23	Length:	1161
Score:	782.00	Matches:	140
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.29%	Mismatches:	0
Query Match:	73.50%	Indels:	0
DB:	9	Gaps:	0
US-09-921-994-2	(1-192) x HUMEBI3X	(1-1161)	
OY	52	HisAspProAlaAlaSerProGlyProCysProLeuGlyGlnLeuProAlaLeuArgTrp	71
DB	13	CATGACCCCGCAGCTTCTCTCGGCCCTTGCTCTTGAGCCAGCTGCCGCCCTGCAGTGG	72

QY	72	LysGluArgAlaProSerSerSerApThrAlaProGlyAlaMetProSerLeuSerVal	91
Db	73	AAGMAAGGGCCCCCAGCAGCTGTGACACTGCCCGGGGTGCAATGCCGAGCCTCTCGTA	132
QY	92	ProAspArgArgGlyLeuLeuLeuAspProAlaAlaCysSerLysLeuHisGlnProArg	111
Db	133	CCCGATCGCCGTGGATTGGCTCCGGACCTGCCCGCTGCTCCAAACTCCACACGCCCCGT	192
QY	112	ValLeuHisCysHisValGlnAlaArgHisGlyCysProGlyProGlnLeuAlaLeuPro	131
Db	193	GTCCTTCATTGCCACGTCACAGGCTCGGACATGGCTGCCCGGGCCACAGCTGGCCCT	252
QY	132	AlaAlaAspAlaAsnValHisGlnLeuHisHisGlyCysProAlaValLeuHisGly	151
Db	253	GCAGCAGACGCCAACGTCACACGAGCTGCACCATCAGGATGTCAGCTTTCTCCATGGC	312
QY	152	SerLeuArgAlaGlnCysHisArgArgProProLeuGlyLeuGlnGlnLeuArgAla	171
Db	313	TCCCTACGTGCTCAATGTCACCGCCGTCACCCCTGGGGCTCCAGCAGCAGCTTCGTGCC	372
QY	172	PheHisAsnArgAlaHisHisGlnAlaArgProSerArgArgAlaProLysProPro	191
Db	373	TTTCATAACAGACACATCATCAAGCCCGACCTCCAGAAGGCGGTGGCCCTTAAGCCCCCT	432
QY	192	Arg	192
Db	433	CGC	435
RESULT	6		
AC005578		43297 bp	DNA linear PRI 01-SEP-1998
LOCUS			
DEFINITION	Homo sapiens chromosome 19, cosmid F20887, complete sequence.		
ACCESSION	AC005578		
VERSION	AC005578.1	GI:3510237	
KEYWORDS	HTG.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 43297)		
AUTHORS	Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Adamson,A.W.,		
	Burkhardt-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stillwagen,S.,		
	Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Garies,J.,		
	Danganan,L., Poundstone,P., Christensen,M., Georgescu,A., Avila,J.,		
	Liu,S., Attix,C., Andreise,T., Trankheim,M., Amico-Keller,G.,		
	Coeffield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,		
	Krommiller,B., Arellano,A., Montgomery,M., Ow,D., Nolan,M.,		
	Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.		
	Sequence analysis of a 2.5 Mb region in 19p13.3		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 43297)		
REFERENCE	Lamerdin,J.E.		
AUTHORS	Direct Submission		
TITLE	Submitted (01-SEP-1998) Joint Genome Institute, Lawrence Livermore		
JOURNAL	National Laboratory, 7000 East Ave., Livermore, CA 94551, USA		
	Map and sequence oriented from p telomere to centromere. Cosmid		
	F20887 overlaps cosmid R33243 to the left by at least 8 kb and		
	overlaps BC84616 (CUT-B-144d21) to the right from bases 41,733 to		
	43,297. Additional chromosome 19 map and sequence information may		
	be obtained at: <a href="http://www-bio.llnl.gov/bbrp/genome/genome.html">http://www-bio.llnl.gov/bbrp/genome/genome.html</a> .		
COMMENT	location/Qualifiers		
	1. .43297		
FEATURES	/organism="Homo sapiens"		
source	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map="19p13.3 between D19S883 and D19S325"		
	/clone="F20887"		
	/cell_line="UV5HL9-5B"		
	/clone_lib="LL19NC02 F chromosome 19-specific cosmid		
	library"		
	/note="Cosmid library constructed at LLNL from flow-sorted		
	chromosomes from human-hamster hybrid UV5HL9-5B, which		

CDS  
carries chromosome 19 as its only human chromosome."  
join(<177..283,457..644,1440..1477,2941..3048,5246..5374,  
8654..8705,8927..9078,9226..9936,12267..12434,  
16803..16868,17104..17181)  
/note="Hypothetical human protein most similar to PID  
91429314 and several ankyrin-like proteins"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F20887\_1, partial CDS"  
/protein\_id="AAC33490.1"  
/db\_xref="GI:3510240"  
/translation="SGATPLIIAQMCHTDLCRLILQQGAANDDLQGR TALMLACE  
GASPTVEVLLOGAQPGITDALGODAAHYGALAGDKLILHLLOEAAQSPSPALTE  
DDSGEASSQDDRDAEEIVRLROERGRLQKIRGLEQHKERQQRQLSHCCSQVE  
LQQLVERQEEKESLGREVESLOSRLSLENERENTSYDVTTLQDEEGPQGLLRPP  
GAEVLLSRQLSPSAQEHLASLOQAVAVLTRNQELMEKVQILENFEKDFQMEVEALA  
EVIPLALYDSLRAEFQOLRQHAELQALRQETREVPREEGACGSESVAGATATKN  
GPTMELNGSVAPETRYNGAETIDEAAGDETMARTMEAATGAATGABATGAKVT  
ETKPTGAEVREMETTEEDANMETKPTGAQATDTEETTGVEAMGVEATKTRAEAEQAY  
GVGAGQAEPPVTGTTNMEATGSRARQASRAQALDKAKEKDKKEAARHSSVVALYRS  
AAEAQLATAEQQLRGLRTEARARQASRAQALDKAKEKDKKEAARHSSVVALYRS  
HLLYAIQGMDEDVQRILSQILQMQLQAQGR"  
318..346  
repeat\_region  
/rpt\_family="MIR"  
457..644  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 93.000"  
repeat\_region  
complement(1043..1258)  
/rpt\_family="AluSg"  
1440..1477  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: good, score: 68.000"  
repeat\_region  
complement(1785..2083)  
/rpt\_family="AluSx"  
complement(2124..2421)  
/rpt\_family="AluSx"  
2734..2814  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 80.000"  
repeat\_region  
2941..3048  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 96.000"  
repeat\_region  
3114..3264  
repeat\_region  
/rpt\_family="LINE2"  
complement(3505..3808)  
/rpt\_family="AluSg"  
4033..4330  
repeat\_region  
/rpt\_family="AluSx"  
4495..4537  
repeat\_region  
/rpt\_family="LINE2"  
4550..4587  
repeat\_region  
/rpt\_family="LINE2"  
4595..4894  
repeat\_region  
/rpt\_family="AluY"  
4968..5001  
repeat\_region  
/rpt\_family="LINE2"  
5246..5374  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 99.000"  
repeat\_region  
5527..5817  
repeat\_region  
/rpt\_family="AluSg"  
complement(5896..6085)  
/rpt\_family="GGAA)n"  
complement(6086..6129)  
/rpt\_family="GGAA)n"  
complement(6130..6323)  
/rpt\_family="AluSg/x"  
complement(6326..6608)  
/rpt\_family="AluSg"  
6700..6833  
repeat\_region  
/rpt\_family="(CA)n"  
complement(6845..7093)  
/rpt\_family="AluJb"

repeat\_region  
7249..7545  
/rpt\_family="AluSx"  
repeat\_region  
7791..7879  
/rpt\_family="L1"  
repeat\_region  
7880..8157  
/rpt\_family="AluSx"  
8161..8464  
/rpt\_family="AluJo"  
8654..8705  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 91.000"  
8960..9087  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 100.000"  
9226..10839  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 92.000"  
complement(10973..11275)  
/rpt\_family="AluSg"  
complement(11334..11425)  
/rpt\_family="(GGAA)n"  
complement(11426..11678)  
/rpt\_family="AluJo"  
11680..11981  
repeat\_region  
/rpt\_family="AluSx"  
11982..12179  
repeat\_region  
/rpt\_family="AluJb"  
12267..12434  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 2, quality: excellent, score: 100.000"  
complement(12536..12569)  
/rpt\_family="MIR"  
complement(12653..12945)  
/rpt\_family="AluSx"  
complement(12957..13135)  
/rpt\_family="MIR"  
complement(13229..13530)  
/rpt\_family="AluY"  
complement(13580..13751)  
/rpt\_family="AluJ"  
complement(13755..14043)  
/rpt\_family="AluSx"  
complement(14074..14203)  
/rpt\_family="AluJ"  
complement(14214..14318)  
/rpt\_family="MIR"  
14358..14430  
repeat\_region  
/rpt\_family="MIR"  
complement(14465..14667)  
/rpt\_family="MER20"  
complement(14697..14731)  
/rpt\_family="LINE2"  
14783..15083  
repeat\_region  
/rpt\_family="AluSx"  
15087..15137  
repeat\_region  
/rpt\_family="MER58B"  
15178..15228  
repeat\_region  
/rpt\_family="MER58B"  
complement(15624..15908)  
/rpt\_family="AluSg"  
complement(15919..16042)  
/rpt\_family="AluJo/FRAM"  
16045..16077  
repeat\_region  
/rpt\_family="(TA)n"  
complement(16078..16246)  
/rpt\_family="AluSx"  
complement(16251..16539)  
/rpt\_family="AluSc"  
complement(16540..16675)  
/rpt\_family="AluSx"  
16803..16868  
misc\_feature  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 2, quality: excellent, score: 89.000"

```
repeat_region      16896. .16955
                    /rpt_family="MIR"
misc_feature        17104. .17181
                    /note="predicted exon, program: grail2exons_human_1.3,
                    frame: 0, quality: excellent, score: 90.000"
repeat_region      complement(17628. .17925)
repeat_region      /rpt_family="AluY"
repeat_region      complement(17926. .18223)
repeat_region      /rpt_family="AluSx"
repeat_region      18327. .18613
repeat_region      /rpt_family="AluJo"
repeat_region      complement(18910. .19211)
```

Alignment Scores:

Pred. No.:	9.75e-06	Length:	43297
Score:	374.50	Matches:	87
Percent Similarity:	61.64%	Conservative:	11
Best Local Similarity:	54.72%	Mismatches:	32
Query Match:	35.20%	Indels:	29
DB:	9	Gaps:	8

US-09-921-994-2 (1-192) x AC005578 (1-43297)

```
QY 53 AsProAlaAlaSerPro-----GlyProCysProLeuGlyGlnLeuProAlaLeu 69
      ::|||::: ::||| ||||| |||||
Db 25599 GAACCATCAGGGGCAAGCGTGTGGGGCCAGCCCT-----GCCCGCGG--- 25646

QY 70 ArgTrp-----LysGluArgAlaProSerSerSerAspThrAlaProGlyAla 85
      ||| ::||| ||| |||
Db 25647 GGGTGGCAGCGGCCACCAAGACCCCGGGGATCCCATCCGCTGCCCTCCTGT 25706

QY 86 --MetProSerLeuSerValProAsp-----ArgArgGlyLeuLeuAspProAla 103
      ::||| ||| ||||| ||||| |||||
Db 25707 TCCTGCCTCTCTTGGGGTCCCGGGTCAAGGCCCGGACAGTAGAGGTGCTGAGGCCACAG 25766

QY 103 lAcysSerLys--LeuHisGlnProArgValLeuHisCysHisValGlnAlaArgHis 122
      || ::|| ::|||
Db 25767 CACTCCCTGAGGGCGCTCAGCGAGCCCCAC-----CCTGTGCAGGCTCGGCATG 25814

QY 122 lYcysProGlyProGlnLeuAlaLeuProAlaAlaAspAlaAsnValHisGlnLeuHis 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25815 GCTGCCCGGGGCCACAGCTGGCCCTGCTGACAGACGCCAACGTCACACAGCTGCACC 25874

QY 142 lHisGlyCysProAlaValLeuHisGlySerLeuArgAlaGlnCysHisArgArgProp 162
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25875 ATCACGGATGTCCAGCTGTCTCCATGGCTCCCTACGTCTCAATGTCTACCGCGCTCCAC 25934

QY 162 roLeuGlyLeuGlnGlnLeuArgAlaPheHisAsnArgAlaHisHis----- 178
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25935 CCCTGGGGCTCCAGCAGCAGCTTCTGCTTTCATACAGACACATCAGTAGTGGGG 25994

QY 179 -----GlnAlaArgProSerArgArgArgAlaProLysPro 190
      ||||| |||||
Db 25995 CCGCAGTGGGGGGGGGGGGGCTGCCGTCTCCTCCAGCTCCCGCCACCCC 26047
```

RESULT 7  
AX392987 445 bp DNA linear PAT 23-MAR-2002  
LOCUS Sequence 29 from Patent WO0214358.  
DEFINITION AX392987  
ACCESSION AX392987  
VERSION AX392987.1 GI:19701034  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Edmonds,B.T., Micanovic,R., Ou,W., Su,E.W., Tschang,S.H. and Wang,H.  
TITLE Novel secreted proteins and their uses  
JOURNAL Patent: WO 0214358-A 29 21-FEB-2002;  
ELI LILLY AND COMPANY (US)  
FEATURES location/Qualifiers

```
source      1. .445
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
CDS         25. .333
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD29041.1"
            /db_xref="GI:19701035"
            /translation="MTPQLLLALVLMASCPGSGRKKPPAALTLPVQGRASRYPIAV
            DCSWTLPPAPNSTSPVSFITATYRSESLGGLRDTGLPGEPTLALRALGSGKLETPTS"
```

mat\_peptide 76. .330  
BASE COUNT 124 a 138 c 115 g 68 t  
ORIGIN

Alignment Scores:

Pred. No.:	2.18e-07	Length:	445
Score:	373.00	Matches:	69
Percent Similarity:	92.11%	Conservative:	1
Best Local Similarity:	90.79%	Mismatches:	6
Query Match:	35.06%	Indels:	0
DB:	6	Gaps:	0

US-09-921-994-2 (1-192) x AX392987 (1-445)

```
QY 52 HisAspProAlaAlaSerProGlyProCysProLeuGlyGlnLeuProAlaLeuArgTrp 71
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24 CATGACCCCGCAGCTTCTCTGCCCCCTGTCTCTGGGCCACAGCTGCCCGCCCTGCAGTGG 83

QY 72 LysGluArgAlaProSerSerSerAspThrAlaProGlyAlaMetProSerLeuSerVal 91
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 AAGGAAAGGGCCCCCAGCAGCTCTGACACTGCCCGGGGTGCAATGCCGAGCCTCTCGGTA 143

QY 92 ProAspArgArgGlyLeuLeuLeuAspProAlaAlaCysSerLysLeuHisGlnProArg 111
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 CCCGATCGCCGCTGATGCTCTCTGAGACCCCTGCCGCTGCTCCAACCTCCACAGCCCCGT 203

QY 112 ValLeuHisCysHisValGlnAlaArgHisGlyCysProGlyProGln 127
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 204 GTCTTCAATTGCCACGTACAGCTCGGAGAGCCTGGAAGGGGGCCTCAG 251
```

RESULT 8  
AC027272/c 179607 bp DNA linear HTG 01-SEP-2000  
LOCUS Homo sapiens chromosome 19 clone RP11-486P13, WORKING DRAFT  
DEFINITION AC027272  
SEQUENCE, 18 unordered pieces.  
AC027272 GI:9838224  
VERSION AC027272.4  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
TITLE 1 (bases 1 to 179607)  
JOURNAL Waterston,R.H.  
AUTHORS The sequence of Homo sapiens clone  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 179607)  
JOURNAL Waterston,R.H.  
TITLE Direct Submission  
AUTHORS Submitted (28-MAR-2000) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
COMMENT On Aug 17, 2000 this sequence version replaced gi:8570236.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site:http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NH0486P13  
----- Summary Statistics -----



FEATURES	source
Sequencing vector: M13; 95%	
Sequencing vector: plasmid; 0%	
Chemistry: Dye-primer ET; 95% of reads	
Chemistry: Dye-terminator Big Dye; 0% of reads	
Assembly program: Phrap; version 0.990319	
Consensus quality: 170867 bases at least Q40	
Consensus quality: 174124 bases at least Q30	
Consensus quality: 176036 bases at least Q20	
Insert size: 188000; agarose-fp	
Insert size: 179350; sum-of-contigs	
Quality coverage: 4.55 in Q20 bases; agarose-fp	
Quality coverage: 4.80 in Q20 bases; sum-of-contigs	
-----	
* NOTE: This is a 'working draft' sequence. It currently * consists of 18 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.	
1	1338: contig of 1338 bp in length
1339	1438: gap of unknown length
1439	3330: contig of 1892 bp in length
3331	3430: gap of unknown length
3431	5429: contig of 1999 bp in length
5430	5529: gap of unknown length
5530	8107: contig of 2578 bp in length
8108	8207: gap of unknown length
8208	11180: contig of 2973 bp in length
11181	11280: gap of unknown length
11281	14039: contig of 2759 bp in length
14040	14139: gap of unknown length
14140	17230: contig of 3091 bp in length
17231	17330: gap of unknown length
17331	21334: contig of 4004 bp in length
21335	21434: gap of unknown length
21435	26877: contig of 5443 bp in length
26878	26977: gap of unknown length
26978	31453: contig of 4476 bp in length
31454	31553: gap of unknown length
31554	36515: contig of 4962 bp in length
36516	36615: gap of unknown length
36616	46678: contig of 10063 bp in length
46679	46778: gap of unknown length
46779	57999: contig of 11221 bp in length
58000	58099: gap of unknown length
58100	74742: contig of 16643 bp in length
74743	74842: gap of unknown length
74843	92504: contig of 17662 bp in length
92505	92604: gap of unknown length
92605	119545: contig of 26941 bp in length
119546	119645: gap of unknown length
119646	147254: contig of 27609 bp in length
147255	147354: gap of unknown length
147355	179607: contig of 32253 bp in length.
Location/Qualifiers	
1. .179607	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="19"	
/clone="RP11-486P13"	
1. .1338	
/note="assembly_name:Contig8"	
1439. .3330	
/note="assembly_name:Contig9"	
3431. .5429	
/note="assembly_name:Contig10"	
5530. .8107	
/note="assembly_name:Contig11"	
8208. .11180	
/note="assembly_name:Contig12"	
11281. .14039	

```

misc_feature      /note="assembly_name:Contig13"  

                  14140. .17230  

                  /note="assembly_name:Contig14"  

misc_feature      17331. .21334  

                  /note="assembly_name:Contig15  

                  clone_end:SP6  

                  vector_side:left"  

misc_feature      21435. .26877  

                  /note="assembly_name:Contig16"  

                  26978. .31453  

                  /note="assembly_name:Contig17"  

misc_feature      31554. .36515  

                  /note="assembly_name:Contig18"  

                  36616. .46678  

                  /note="assembly_name:Contig19"  

misc_feature      46779. .57999  

                  /note="assembly_name:Contig20"  

                  58100. .74742  

                  /note="assembly_name:Contig21"  

misc_feature      74843. .92504  

                  /note="assembly_name:Contig22"  

                  92605. .119545  

                  /note="assembly_name:Contig23"  

                  119646. .147254  

                  /note="assembly_name:Contig24"  

misc_feature      147355. .179607  

                  /note="assembly_name:Contig25"  

BASE COUNT      41107 a 48085 c 48884 g 39812 t 1719 others  

ORIGIN  

Alignment Scores:  

Pred. No.:      0.000257      Length:      179607  

Score:          351.50      Matches:      89  

Percent Similarity: 51.05%      Conservative: 8  

Best Local Similarity: 46.84%      Mismatches:  35  

Query Match:    33.04%      Indels:      60  

DB:            2      Gaps:      7  

US-09-921-994-2 (1-192) x AC027272 (1-179607)  

QY      12  TTPALASerCysPro--ProArgSerGlyArgLysGlyProProAlaAlaLeuThrLeu 30  

         |||      |||||      |||||      |||||      |||||  

Db 158705 TGGGGCCCATGCCCCCTTGCCCCCGGGGGTGGGCACGCGCACCCACCACAGG----- 158659  

QY      31  ProArgValGlnCysARgAlaSerArgTyrProIleAlaValAspCysSerTrpThrLeu 50  

         -----  

Db 158658 -----  

QY      51  ProHisAspProAlaAlaSerProGlyProCysProLeuGlnLeuProAlaLeuArg 70  

         |||      |||||      |||      |||      :::|||||  

Db 158652 CCCCCGGATCCCCCATCCGCTGCCCCCTCT-----GTTCTGCTCTCTCT 158608  

QY      71  TTPlySGluArgAlaProSerSerSerAspThrAlaProGlyAlaMetProSerLeuSer 90  

         |||      -----  

Db 158607 TGG-----  

QY      91  ValProAspArgArgGly-LeuLeuLeuAspProAlaAlaCysSerLysLeuHisGlnPr 110  

         ::|      |||      |||      |||      |||      |||  

Db 158577 GGTGCTGAGGCGCACAGGCACTC-----CCTGAGCGCTTCAGCGAGCCCCACCTCT-- 158529  

QY      110  oArgValLeuHisCysHisValGlnAlaArgHisGlnCysProGlyProGlnLeuAlaLe 130  

         |||||      |||||      |||||      |||||      |||||      ||  

Db 158528 -----GTGACGGCTCGGCATGGCTGCCCCGGGCCACAGCTGGC-CT 158489  

QY      130  uProAlaAlaAspAlaAsnValHisGlnLeuHisHisHisGlnCysProAlaValLeuHi 150  

         |||||      |||||      |||||      |||||      |||||      |||||  

Db 158488 GCCTGCAGCAGACGCCAACGTCACACAGCTGCACCATCACGGATGTCCAGCTGTTCTCCA 158429  

QY      150  sGlySerLeuArgAlaGlnCysHisArgArgProProLeuGlnGlnGlnLeuAr 170  

         |||||      |||||      |||||      |||||      |||||      |||||  

Db 158428 TGGCTCCCTACGTGCTCAATGTCACCGCGCTCCACCCCTGGGGCTCCAGCAGCACTTCG 158369  

QY      170  gAlaPheHisAsnArgAlaHisHis-----GlnAlaAr 181

```





*	145606	145705:	gap of	unknown	length
*	145706	148555:	contig	of 2850 bp	in length
*	148556	148655:	gap of	unknown	length
*	148656	151425:	contig	of 2770 bp	in length
*	151426	151525:	gap of	unknown	length
*	151526	158286:	contig	of 6761 bp	in length
*	158287	158386:	gap of	unknown	length
*	158387	160638:	contig	of 2252 bp	in length
*	160639	160738:	gap of	unknown	length
*	160739	173388:	contig	of 12650 bp	in length
*	173389	173488:	gap of	unknown	length
*	173489	179077:	contig	of 5589 bp	in length
*	179078	179177:	gap of	unknown	length
*	179178	184518:	contig	of 5341 bp	in length
*	184519	184618:	gap of	unknown	length
*	184619	189175:	contig	of 4557 bp	in length
*	189176	189275:	gap of	unknown	length
*	189276	191154:	contig	of 1879 bp	in length
*	191155	191254:	gap of	unknown	length
*	191255	192103:	contig	of 849 bp	in length
*	192104	192203:	gap of	unknown	length
*	192204	193355:	contig	of 1152 bp	in length.

FEATURES	Location/Qualifiers
source	1. .193355
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/clone="RP23-264G1"
	/clone_lib="RPCI mouse BAC library 23"
BASE COUNT	49405 a 44453 c 45086 g 52104 t 2307 others
ORIGIN	

Alignment Scores:		
Pred. No.:	360	193355
Score:	193.50	73
Percent Similarity:	40.35%	19
Best Local Similarity:	32.02%	74
Query Match:	18.19%	62
DB:	2	8
		Gaps:

US-09-921-994-2 (1-192) x AC073737 (1-193355)

QY	24	ProProAla-----AlaLeuThrLeuProArgValGlnCysArgAlaSerArgTyrPro	41
Db	112289	CCTCCAGCTCTCGTGCTCTAAGCCAGGCCAGAGTCATGCCATGTCTTCGGTATCCC	112230
QY	42	IleAlaValAspCysSerTrpThr-LeuProHisAspProAlaIaaserProGlyProCy	61
		:::	
Db	112229	GTCGCCGTGACTGCTCTCGACTCTCTCCAGGCTCCCA--ACTCCACCAGATCCACG	112173
QY	61	sProLeu-GlyGlnLeuProAlaLeuArgTrpLys-----GluArgAlaProSer-Ser	78
Db	112172	TTCCTTCATTGCCACTTACAGGTCAGAAATGAAAGGTTTAGGAGAAGCCAGATATCA	112113
QY	79	SeraSpThr-----AlaProGlyAlaMetProSerLeu	89
		:::	
Db	112112	AGAGATTCCAATGTCTCATCCACTCAGGGTAGAGCAGCCTAGGCCCTTGGGAACCTG	112053
QY	90	SerValProAspArgArg-----GlyLeuLeuLeuAspPro	101
		:::	:::
Db	112052	CTTCTTAAGCAACAGACAAGATGTGGTTGTGGTATTCTTAGGAGTCTTACAGATCCA	111993
QY	102	AlaAlaCysSer-----	105
		:::	
Db	111992	CAGTCCCCAAGTCCCCCTGAGCATATCACCCCTGAGCGGTGACCCACAGTCAGCTGTGGG	111933
QY	106	-----LysLeuHisGlnProArgValLeuHis-----CysHisVal	117
Db	111932	TGAGCTGGGGCCTCACGGGCACCCAGGTCTTCCTTGAGATTCCACCATGTGCTTCCCATG	111873
QY	118	GlnAlaArgHisGlyCysProGlyProGlnLeuAlaLeuProAlaIaAspAlaAsnVal	137
Db	111872	CAGGCTCGGTGTGGCCACCCAGCAGCAGACAGCCAGCCCTGCTTCAACAGGAGCCCCCAGGC	111813

QY	138	HISGlnLeuHISHisGlyCysProAlaValLeuHISGlySerLeuArgAlaGlnCys	157
		:::	:::
Db	111812	CTCCCGATGCACCATCCCGACGTGCACACTGTCTTCACAGGTGCCCCATCATGCTAAATGT	111753
QY	158	HISArgArgProProLeuGlyLeuGlnGlnGlnLeuArgAlaPheHISAsnArgAlaHis	177
			:::
Db	111752	CACCTGCAGTGCACCCAGCGCGCCAGCAGCAGCCCTGTAAGCCTTGTGTGGCTGAGCGAAT	111693
QY	178	-----HISGlnAlaArgProSerArg	184
Db	111692	CAGTGAAGTGTGTCTCTCCCTGCCCCCTCTCCACCTGCTCCATCCCGTGGCTCCCTCCCTGC	111633
QY	185	ArgArgAlaProLysPro	190
Db	111632	CGCCGCCCCCTTACCCCC	111615

```

RESULT 12
AC118074/c      176359 bp      DNA      linear      HTG 18-JUL-2002
LOCUS          Rattus norvegicus clone CH230-47J17, *** SEQUENCING IN PROGRESS
DEFINITION
***, 69 unordered pieces.

```

ACCESSION	AC118074
VERSION	AC118074.4
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus

ORGANISM  
Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 176359)  
Muzny, D. M., Adams, C., Adio-Oduola, B., Ali-osman, F. R., Allen, C.,

REFERENCE	AUTHORS
1 (bases 1 to 176359)	Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbacia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Koryah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H., Lozador,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sulton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
TITLE	Direct Submission
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 176359)



AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (13-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 3 (bases 1 to 176359)  
AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (18-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
COMMENT On Jul 14, 2002 this sequence version replaced gi:20177536.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GUEB  
Center clone name: CH230-47J17  
----- Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 115372 bases at least Q40  
Consensus quality: 121114 bases at least Q30  
Consensus quality: 124422 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 69 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1219: contig of 1219 bp in length  
\* 1220 1319: gap of unknown length  
\* 1320 2744: contig of 1425 bp in length  
\* 2745 2844: gap of unknown length  
\* 2845 4066: contig of 1222 bp in length  
\* 4067 4166: gap of unknown length  
\* 4167 5794: contig of 1628 bp in length  
\* 5795 5894: gap of unknown length  
\* 5895 6936: contig of 1042 bp in length  
\* 6937 7036: gap of unknown length  
\* 7037 8767: contig of 1731 bp in length  
\* 8768 8867: gap of unknown length  
\* 8868 10546: contig of 1679 bp in length  
\* 10547 10646: gap of unknown length  
\* 10647 12271: contig of 1625 bp in length  
\* 12272 12371: gap of unknown length  
\* 12372 13483: contig of 1112 bp in length  
\* 13484 13583: gap of unknown length  
\* 13584 14951: contig of 1368 bp in length  
\* 14952 15051: gap of unknown length  
\* 15052 16270: contig of 1219 bp in length  
\* 16271 16370: gap of unknown length  
\* 16371 17804: contig of 1434 bp in length  
\* 17805 17904: gap of unknown length  
\* 17905 19394: contig of 1490 bp in length  
\* 19395 19494: gap of unknown length  
\* 19495 20712: contig of 1218 bp in length  
\* 20713 20812: gap of unknown length  
\* 20813 22799: contig of 1987 bp in length  
\* 22800 22899: gap of unknown length  
\* 22900 24177: contig of 1278 bp in length  
\* 24178 24277: gap of unknown length  
\* 24278 25821: contig of 1544 bp in length  
\* 25822 25921: gap of unknown length  
\* 25922 27784: contig of 1863 bp in length

\* 27785 27885: gap of unknown length  
\* 27885 28934: contig of 1050 bp in length  
\* 28935 29034: gap of unknown length  
\* 29035 30354: contig of 1320 bp in length  
\* 30355 30454: gap of unknown length  
\* 30455 31701: contig of 1247 bp in length  
\* 31702 31801: gap of unknown length  
\* 31802 33337: contig of 1536 bp in length  
\* 33338 33437: gap of unknown length  
\* 33438 34728: contig of 1291 bp in length  
\* 34729 34828: gap of unknown length  
\* 34829 36591: contig of 1763 bp in length  
\* 36592 36691: gap of unknown length  
\* 36692 38003: contig of 1312 bp in length  
\* 38004 38103: gap of unknown length  
\* 38104 40127: contig of 2024 bp in length  
\* 40128 40227: gap of unknown length  
\* 40228 41691: contig of 1464 bp in length  
\* 41692 41791: gap of unknown length  
\* 41792 43535: contig of 1744 bp in length  
\* 43536 43635: gap of unknown length  
\* 43636 45466: contig of 1831 bp in length  
\* 45467 45566: gap of unknown length  
\* 45567 47021: contig of 1455 bp in length  
\* 47022 47121: gap of unknown length  
\* 47122 48801: contig of 1680 bp in length  
\* 48802 48901: gap of unknown length  
\* 48902 49935: contig of 1034 bp in length  
\* 49936 50035: gap of unknown length  
\* 50036 51484: contig of 1449 bp in length  
\* 51485 51584: gap of unknown length  
\* 51585 53366: contig of 1782 bp in length  
\* 53367 53466: gap of unknown length  
\* 53467 55706: contig of 2240 bp in length  
\* 55707 55806: gap of unknown length  
\* 55807 57484: contig of 1678 bp in length  
\* 57485 57584: gap of unknown length  
\* 57585 59524: contig of 1940 bp in length  
\* 59525 59624: gap of unknown length  
\* 59625 61504: gap of unknown length  
\* 61505 61604: gap of unknown length  
\* 61605 63038: contig of 1434 bp in length  
\* 63039 63138: gap of unknown length  
\* 63139 65148: contig of 2010 bp in length  
\* 65149 65248: gap of unknown length  
\* 65249 68165: contig of 2917 bp in length  
\* 68166 68265: gap of unknown length  
\* 68266 70437: contig of 2172 bp in length  
\* 70438 70537: gap of unknown length  
\* 70538 72778: contig of 2241 bp in length  
\* 72779 72878: gap of unknown length  
\* 72879 74865: contig of 1987 bp in length  
\* 74866 74965: gap of unknown length  
\* 74966 78152: contig of 3187 bp in length  
\* 78153 78252: gap of unknown length  
\* 78253 80830: contig of 2578 bp in length  
\* 80831 80930: gap of unknown length  
\* 80931 83775: contig of 2845 bp in length  
\* 83776 83875: gap of unknown length  
\* 83876 86764: contig of 2889 bp in length  
\* 86765 86864: gap of unknown length  
\* 86865 90526: contig of 3662 bp in length  
\* 90527 90626: gap of unknown length  
\* 90627 92648: contig of 2022 bp in length  
\* 92649 92748: gap of unknown length  
\* 92749 96414: contig of 3666 bp in length  
\* 96415 96514: gap of unknown length  
\* 96515 99630: contig of 3116 bp in length  
\* 99631 99730: gap of unknown length  
\* 99731 102795: contig of 3065 bp in length  
\* 102796 102895: gap of unknown length  
\* 102896 105478: contig of 2583 bp in length



CDS complement(1806. .2969)  
/gene="acda"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="acyl-CoA dehydrogenase"  
/protein\_id="AAM39565.1"  
/db\_xref="GI:21111212"  
/translation="MHVPSLNFELGEDIDLRESVAAFASHHIAPLAAADHDNVEPPA  
QLWRLFEQGLGLTVEEAYGSGMGYLAHVAMEEISRAGAGISYGAHSNLCNQ  
LRKNATHEQOKRYLPKICTGEHVALAMSEAGSDVSMKLRAEARGDRFVLNSKM  
WITNGPDADVLVYAKTDPGARGITAFIVEKGMPESTAKLDKIGMRGSRNCELY  
FTDCEVPAENVLTGTVGVRVMSGLDFERNVLGGPGLGMAAADVLPYHERKQF  
GEPITGFMQAKLADMYVGLMACRAVYAVARACDAGRTTRQDAAGAILYAAEKATW  
LTGQAIQVLGNGYINIEYPTGRIMRDAKLYEIGAGTSEIRMLIGRELFEETA"  
3098. .3715  
/gene="XCC0247"  
3098. .3715  
/gene="XCC0247"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="transcriptional regulator acrr family"  
/protein\_id="AAM39566.1"  
/db\_xref="GI:21111213"  
/translation="MAYRRSALMEERLAGNRERILHAARALIAEGGYRNAPYTAVAAA  
AGVSTGOIYRHFPSKADLEFEVLAQVQHEMGILRAISAEPVSAEERLRAIAAEVRR  
ALAGPALAYAFIAEPVEGEVDAERIRGRRLFGEYRQLAEGVAAGEFPEQSLDAAA  
CIVGAFTEALVGPVAPSRHDPQGEQLVEAIGCFCLRAAGARGPG"  
3774. .4175  
/gene="XCC0248"  
3774. .4175  
/gene="XCC0248"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM39567.1"  
/db\_xref="GI:21111214"  
/translation="MHSPCRSRAMAVLNPLALTERLRVPDRITLSPSAMAGLLDSQ  
OELAELAGVHRNSLRVHPESPVRQDLRLNSRLLVAMAQIQPERQLVHLKNTPIPA  
FEFATLLEVVKOGRTDALTLYLRTVAGAAG"  
4172. .4708  
/gene="XCC0249"  
4172. .4708  
/gene="XCC0249"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM39568.1"  
/db\_xref="GI:21111215"  
/translation="MKLSAPACITLYRAFTPRMAAEPLSGAGARSGGRFNREGOPAL  
YLSLHLDTAAAEYAOAAPFLPPTLVSYRADLPALADVRDLASWDALWADWTDWARK  
ALVNKYEPVSWVLGDLREAIIPGVIFPMSAQPGGVNALFLDMLQPEQVLHVLDGGR  
LPRDGRSWADPAAERPE"  
complement(5006. .5341)  
/gene="XCC0250"  
complement(5006. .5341)  
/gene="XCC0250"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM39569.1"  
/db\_xref="GI:21111216"  
/translation="MSVYLLALLNFSGFALCLAMEKHQHEVGRGALGTRSRQLRAL  
GWLILLVTFALAVSAQGWIGSVLWLTSLAGAAVLSWLLPYRRGLILPAIAAPV

gene ACMAFALLH"  
complement(5338. .6954)  
/gene="XCC0251"  
complement(5338. .6954)  
/gene="XCC0251"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM39570.1"  
/db\_xref="GI:21111217"  
/translation="MKOGEFROSMAWLHWTGLLVGWVLLIFMGSTASYRDELSRMW  
RPELPTTVSSAVAMRSABERYLQTHAPDAQSWNTLLPDARTPVYTMQNPAPPGKT  
LSRRELYGNAIIDPATGNPISARDTLGDFEYRLHFDLHYLPVWARYIVGFCAMFEL  
VAIISGVITHKKIKFDEFTERPGKGLRSLWDFHNVSATVLPYHAMITYTGIVTLMEM  
YLPWGIIKAQYPDNEMREYEEESANRVADTRTAAGTPARMRPLEFFVARARSDWRGDDVG  
TVAVSLPNDAAHAAGVTVQRADDLNDAPSILYDAVSGORLYSGAPGASQTRGVNVG  
LHIAHFAGWMRALFFGSGLLGLMVAASGVVMVAVKERPKHAKAGRIGFGLRLVDALN  
IGTVAGLPIAFPAEFFWANRLVPVALLEARAMEAHLFFAAMGALLAAPVPRRAMSW  
QLYLGAGLFALVPLNALTTDLHLGVTVPAGQWALAGVDLVCGLGVCLGIAGWRIDH  
WKAPOSAAARKARAQATPAPONALEPVQEGA"  
complement(6951. .7262)  
/gene="XCC0252"  
complement(6951. .7262)  
/gene="XCC0252"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAM39571.1"  
/db\_xref="GI:21111218"  
/translation="MSMASPLPAPWFKRPWLGVLRVAAIGGGYALASATNLLAVV  
LPMRSEAVLTSMLVGIVVCACAPLWAFATASVLRAMVIGVPAVAMFAAAMWLQRGAA  
A"  
complement(7281. .7928)  
/gene="XCC0253"  
complement(7281. .7928)  
/gene="XCC0253"  
/note="putative; ORF located using Glimmer/Genemark"

Alignment Scores:  
Pred. No.: 1.52e+03. Length: 10770  
Score: 149.50 Matches: 71  
Percent Similarity: 37.77% Conservative: 17  
Best Local Similarity: 30.47% Mismatches: 84  
Query Match: 14.05% Indels: 62  
DB: 1 Gaps: 13

US-09-921-994-2 (1-192) x AE012121 (1-10770)

QY 2 ThrProGlnLeuLeuAlaLeuValLeuTrpAlaSerCysProProArgSerGlyArg 21  
DB 5023 ACGCCATCCCGCGCAGCAGCAGTGCAGCAATG-----CTGCCGCGCAAGATCAGGCCG 5076

QY 22 LysGlyProProAla-----AlaLeuThrLeuProArgValGlnCysArgAlaSerArg 39  
DB 5077 GCCCGTACGGCAGCAACACAGCAGCAGCGCGCCTGCGC-----TCAGCG 5127

QY 40 TyrProIleAlaValAspCysSerTrpThrLeuProHisAspProAlaAlaSerProGly 59  
DB 5128 TGCCCAAGCCACAGTACGCTG-----CGATGCCCCAGCCTTGCGCCGACA 5172

QY 60 ProCysProLeuGlnLeuProAlaLeuArgTrpLysGluArgAla-----ProSer 77  
DB 5173 CCG-----CCAGCGCAAAAGGTGACCAACAGCAGCAGCAGCATCCACAGCG 5214

QY 78 SerSerAspThrAlaProGlyAlaMetProSerLeuSerValProAspArgArgGlyLeu 97  
DB 5215 CGCGCAACTGACGTGAGCGGGTCATGCCACGCGCGCCGACGACGACCTGCTGCTGT 5274

QY 98 LeuLeuAspProAla----- 102

Db 5275 TTTCATCGCCAGGACAAAGCGCGGAAGCCGGAAATTCAGGCCACACAGCACTA 5334
QY 103 ----AlaCysSerLysLeuHisGlnProArgValLeuHisCysHisValGlnAlaArgHis 121
Db 5335 CGCTCATGCGCCTTCCTGCACAGCAGCGCGTTTGGCGGCGCAGGCGCTTGGCGCG 5394
QY 122 GLy-----CysProGly-----ProGlnLeuAlaLeuProAlaAlaAsp 134
Db 5395 GGCCTTGCGGTGCGCGCGGATTCGGCGCGCTTCACGTGCTGCAGGCGGCCGCGCAT 5454
QY 135 -----AlaAsnValHisGlnLeuHisHisGlyCysPro---Ala 147
Db 5455 GCCCAGGCACAGCCCGCAGGCCACACACAGATCCAGCGCGCCAATGCCCATTTGGCC 5514
QY 148 ValLeuHisGlySerLeuArgAlaGlnCys-----HisArg 159
Db 5515 TGCCGGCAGCGTCAAGCGGAGGTGCGGTGTCAGCGCGCTTGAGCAGCGGCACCA 5574
QY 160 ArgProProLeuGlyLeuGlnGlnLeuArgAlaPheHisAsnArgAlaHisGln 179
Db 5575 CGCGAACCAAGCCTGCGCCCGCAGGTACAGTCCACGCTCCACATCG-----CACGCC 5624
QY 180 AlaArgProSerArgArgArgAlaProLysProProArg 192
Db 5625 GCGGCCACACGAAGGCGCGAGCAAGGCCGCCACCCAGG 5663
RESULT 14
AC131212 168964 bp DNA linear HTG 18-AUG-2002
LOCUS Homo sapiens clone RP13-554M15, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 23
DEFINITION unordered pieces.
AC131212
AC131212.1 GI:22296917
VERSION HTG; HTGS\_PHASE1.
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 168964)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
HomsI,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Monabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,S., Oguh,M., Okwuonu,G.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Nguyen,A., Nguyen,N.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: HEME
Center clone name: RP13-554M15
Center clone name: Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 149776 bases at least Q40
Consensus quality: 152075 bases at least Q30
Consensus quality: 153483 bases at least Q20
\* NOTE: Estimated insert size may differ from sequence length
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
\* NOTE: This is a 'working draft' sequence. It currently
\* consists of 23 contigs. The true order of the pieces
\* is not known and their order in this sequence record is
\* arbitrary. Gaps between the contigs are represented as
\* runs of N, but the exact sizes of the gaps are unknown.
\* This record will be updated with the finished sequence
\* as soon as it is available and the accession number will
\* be preserved.
1 2409: contig of 2409 bp in length
2410 2509: gap of unknown length
2510 4892: contig of 2383 bp in length
4893 4992: gap of unknown length
4993 7194: contig of 2202 bp in length
7195 7294: gap of unknown length
7295 9447: contig of 2153 bp in length
9448 9547: gap of unknown length
9548 12910: contig of 3363 bp in length
12911 13010: gap of unknown length
13011 16196: contig of 3186 bp in length
16197 16296: gap of unknown length
16297 19559: contig of 3263 bp in length
19560 19659: gap of unknown length
19660 24011: contig of 4352 bp in length
24012 24111: gap of unknown length
24112 28769: contig of 4658 bp in length
28770 28869: gap of unknown length
28870 34551: contig of 5682 bp in length
34552 34651: gap of unknown length
34652 39215: contig of 4564 bp in length
39216 39315: gap of unknown length
39316 43668: contig of 4353 bp in length
43669 43768: gap of unknown length
43769 50767: contig of 6999 bp in length
50768 50867: gap of unknown length
50868 57852: contig of 6985 bp in length
57853 57952: gap of unknown length
57953 64498: contig of 6546 bp in length
64499 71188: gap of unknown length
71189 71288: gap of unknown length



71289 \* 80962: contig of 9674 bp in length  
\* 80963 81062: gap of unknown length  
\* 81063 92068: contig of 11006 bp in length  
\* 92069 92168: gap of unknown length  
\* 92169 105140: contig of 12972 bp in length  
\* 105141 105240: gap of unknown length  
\* 105241 119456: contig of 14216 bp in length  
\* 119457 119556: gap of unknown length  
\* 119557 133120: contig of 13564 bp in length  
\* 133121 133220: gap of unknown length  
\* 133221 150729: contig of 17509 bp in length  
\* 150730 150829: gap of unknown length  
\* 150830 168964: contig of 18135 bp in length.  
Location/Qualifiers  
1. 168964  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP13-554M15"  
BASE COUNT 36032 a 46964 c 48094 g 35152 t 2722 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.69e+04 Length: 168964  
Score: 149.00 Matches: 66  
Percent Similarity: 39.59% Conservative: 12  
Best Local Similarity: 33.50% Mismatches: 77  
Query Match: 14.00% Indels: 42  
DB: 2 Gaps: 10  
US-09-921-994-2 (1-192) x AC131212 (1-168964)  
OY 12 TrpAlaSerCysPro--ProArgSerGlyArgLysGlyPro-ProAlaAlaLeuThrLe 30  
Db 127607 TGGTGGCGGACCCCGTGCCTGCTGCTCCCGGTGGCCGACCGGCGACACCTT----- 127661  
OY 30 uProArgValGlnCysArgAlaSerArgTyrProIleAlaValAspCysSer-TrpThrL 50  
Db 127662 -----GCCGAGGAGAGAGGGCTTGGGGACCCGGGACCCAGCACTGCCCTGTGG---- 127710  
OY 50 euProHisAspProAlaAlaLaserProGlyProCysProLeuGlyGlnLeuProAlaLeuA 70  
Db 127711 --CCTGCGTGAACCTGCTTACGGTAAAGCCACGATACCCAGAGAAAGGGGAGGGCGGCTCA 127768  
OY 70 rGTrpLysGlu-----ArgAlaProSerSerSerAspThrAlaProGlyAlaMetPros 88  
Db 127769 CGCGTGAAGATCACACCCGCTGCACACGGCTGTGAACACACAGACCTCATGCTGCGCT- 127827  
OY 88 erLeuSerValProAspArgGlyLeuLeuLeuAspProAlaAlaCysSerLysLeuH 108  
Db 127828 -----TCAGGGGCGGACGAGGCTGAGTC-----GGGTTCAACAAGCGCCCTGC 127870  
OY 108 lsgLInProArgValLeuHis-----C 115  
Db 127871 CCCAGCCGAGGTGTTCTACCTTTATCTTCTGCTGAGCGGTAGATCTGCCAGGCGATC 127930  
OY 115 yshIsValGlnAlaArgHisGlyCysProGlyProGlnLeuAlaLeuProAlaAlaAspA 135  
Db 127931 TGCACGTG-----CAGGGCACACCACTCCCGGCTTCTGCGACAGATGGAAGCG 127981  
OY 135 laAsnValHisGlnLeuHisHisGlyCys-----ProAlaValLeuHisG 151  
Db 127982 CTCAGTGGGGCTGTGTCACCCCAACCATGTTTCTGCAGACTCCTGCTGTACAGGGGCC 128041  
OY 151 lySerLeuArgAlaGlnCysHisArgArgProProLeuGlyLeuGlnGlnLeuArgA 171  
Db 128042 CCCCAGAGCGCGCCGCGGGGACACAGAGTCTCTCCCTTGCCA-----CGGG 128086  
OY 171 laPheHisAsnArgAlaHisHisGlnAlaArgProSerArgArgArg 186  
Db 128087 CCTTCCCAACACAGTGTCACTTGACCACCGCGGTACGGGTACAGA 128133  
RESULT 15  
AC079031

LOCUS AC079031 186558 bp DNA linear HTG 31-JUL-2002  
DEFINITION Homo sapiens chromosome 12 clone RP11-503G7, WORKING DRAFT  
ACCESSION AC079031  
VERSION AC079031  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vetr,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 186558)  
AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 3 (bases 1 to 186558)  
AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
COMMENT On Jul 29, 2002 this sequence version replaced gi:21908302.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Drafting Center Code: BCM  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information

